INTRODUCTION TO FIELD THEORY

Victor F. Weisskopf

(Lectures given to the experimental physicists at CERN)
Propriété littéraire et scientifique réservée pour tous les pays du monde. Ce document ne peut être reproduit ou traduit en tout ou en partie sans l'autorisation écrite du Directeur général du CERN, titulaire du droit d'auteur. Dans les cas appropriés, et s'il s'agit d'utiliser le document à des fins non commerciales cette autorisation sera volontiers accordée.

Le CERN ne revendique pas la propriété des inventions brevetables et dessins ou modèles susceptibles de dépôt qui pourraient être décrits dans le présent document ; ceux-ci peuvent être librement utilisés par les instituts de recherche, les industriels et autres intéressés. Cependant, le CERN se réserve le droit de s'opposer à toute revendication qu'un usager pourrait faire de la propriété scientifique ou industrielle de toute invention et tout dessin ou modèle décrits dans le présent document.

© Copyright CERN, Genève, 1976

Literary and scientific copyrights reserved in all countries of the world. This report, or any part of it, may not be reprinted or translated without written permission of the copyright holder, the Director-General of CERN. However, permission will be freely granted for appropriate non-commercial use. If any patentable invention or registrable design is described in the report, CERN makes no claim to property rights in it but offers it for the free use of research institutions, manufacturers and others. CERN, however, may oppose any attempt by a user to claim any proprietary or patent rights in such inventions or designs as may be described in the present document.
Introduction to Field Theory

by

Victor F. Weisskopf

(Lectures given to the experimental physicists at CERN)
Introductory note

These lectures, which are now appearing as a report, were conceived as an introductory course for experimental physicists at CERN. The course was of a very informal character. It was often referred to as "Threepenny-Field-Theory". Many discussions took place during and between the lectures. Therefore the notes contain some repetitions, some second attempts to improve the explanation of some difficult points, of which there are so many in this field. Because of the restricted time, many important parts of field theory such as the quantization of the electromagnetic field and the interaction of two quantized fields are omitted.

The notes of the lectures given have been written up in a casual manner, essentially for the benefit of those who attended the course. As so many requests for copies of these lectures have come from outside CERN, it has been decided to make them more generally available. Neither the treatment nor the list of references should be considered in any way as definitive.
It is very often that concepts in field theory are not really new but can be understood already in classical physics. We shall therefore start with a simple classical problem.

1. The Harmonic Oscillator in Classical Physics

The Hamiltonian (i.e. the energy) is given by:

\[ H = \frac{1}{2m} p^2 + \frac{\omega^2 m}{2} q^2 \]

where \( p \) is the linear momentum
\( q \) is the coordinate
\( m \) is the mass
and \( \omega \) is the circular frequency.

We also need the Lagrangian (from which one should really start)

\[ L = \frac{m}{2} \dot{q}^2 - \frac{\omega^2 m}{2} q^2 \]

\[ \uparrow \quad \uparrow \]

kin. energy  potential energy

The Lagrangian is defined by the fact that one gets the equation of motion from it in the following way: one first finds the momentum:

\[ p = \frac{\delta L}{\delta \dot{q}} = m \dot{q} \]

then the equation of motion is:

\[ \frac{d}{dt} p = \frac{\delta L}{\delta q} = -\omega^2 mq \]

From (3) and (4) one obtains:
\( m\ddot{q} = -\omega^2 m\dot{q} \)

(6) \( q = A \cos (\omega t + \delta) \)

Where \( A \) is the amplitude of the oscillator and \( \delta \) is a phase.

One obtains the momentum by substituting (6) into (3):

(7) \( p = m\dot{q} = -m\omega A \sin (\omega t + \delta) \)

And the energy by substituting (6) and (7) into (1):

(8) \( H = \frac{p^2}{2m} \)

One can also express the energy by a dimensionless number \( \xi \) so that:

(9) \( H = \hbar \omega \xi \quad (\text{Def. of } \xi) \)

Where \( \hbar \) is some number with the dimension of an action which at this point has nothing to do with quantum theory; \( \xi \) comes out to be:

(10) \( \xi = \frac{A^2 \omega}{2\hbar} \)

Now the problem is essentially solved; there are two independent constants \( A \) and \( \delta \). But we would like to construct a new quantity of which essentially \( q \) is the real part and \( p \) the imaginary part; with a proper choice of constants:

(11) \( C = \sqrt{\frac{\omega m}{2\hbar}} \quad \frac{1}{\sqrt{2m\omega\hbar}} \quad p \)

Which by (6) and (7) becomes:

(12) \( C = \sqrt{\frac{\omega m}{2\hbar}} A \cos (\omega t + \delta) - i \sin (\omega t + \delta) \)
It is seen that the factor in front is simply \( \sqrt{\zeta} \) (from 10) and the parenthesis is equal to \( e^{-i(\omega t + \delta)} \), so that we obtain the simple result:

\[
(13) \quad C = \sqrt{\zeta} \quad e^{-i(\omega t + \delta)}
\]

This quantity has three advantages:
First, being a complex quantity, it contains information about two separate quantities, \( p \) and \( q \).
Secondly, it has a well-defined time-dependance, determined by the minus sign in the exponential.
Thirdly, its amplitude is simply the square root of the energy in its natural units.

\( C \) is of course still a classical quantity (we shall see later that \( C \) and \( C^* \) correspond to destruction and creation operators).

Now one can express \( p \) and \( q \) in terms of \( C \) and \( C^* \) by adding or subtracting eq. (11) from its complex conjugate:

\[
(14) \quad q = \sqrt{\frac{-\hbar}{2m\omega}} \quad (C + C^*)
\]

\[
(15) \quad p = \frac{1}{i} \sqrt{\frac{\hbar \omega m}{2}} \quad (C - C^*)
\]

and by multiplying (11) with its complex conjugate and comparing to (1), one finds a simple expression for the energy:

\[
(16) \quad H = \hbar \omega CC^* = \hbar \omega, C^2
\]

But since \( CC^* = \xi \) according to (13), one also has:

\[
(17) \quad H = \xi \hbar \omega
\]

as it should be because of the definition of \( \xi \) (9). \( C \) contains all the information about the oscillator in a simple way.
2. The Harmonic Oscillator in Quantum Mechanics

Let us use the simplest scheme: the Schrödinger scheme. We replace $p$, $q$ and $H$ by their operators:

(18) \[ (p)_{op} = \frac{\hbar}{i} \frac{\partial}{\partial q} \quad |\psi'(q,t)\rangle \]

(19) \[ (q)_{op} = q \quad |\psi'(q,t)\rangle \]

(19') \[ (H)_{op} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2} + \frac{\nu^2 m}{2} q^2 \quad |\psi'(q,t)\rangle \]

(i.e. the operator corresponding to $p$ is essentially differentiation of the wave function with respect to the coordinate $q$, and the operator corresponding to $q$ is simply multiplication of the wave function by the coordinate $q$, etc.)

One obtains the wave equation:

(20) \[ \left( -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial q^2} + \frac{\nu^2 m}{2} q^2 \right) \psi = -\frac{\hbar}{i} \frac{\partial \psi'}{\partial t} \]

Here and in the future we shall try to get as much knowledge as we can without explicitly solving the Schrödinger equation. Now the wave function can be written as a product of two parts:

(21) \[ \psi_n = \psi_n(q) \ e^{-\frac{i En}{\hbar}} \]

(the minus sign is due to the arbitrary choice of sign for the right hand side of (19'). The opposite sign would correspond to an inverted time scale.) The problem has reasonable bound solutions only if:

(22) \[ E_n = \hbar \omega (n + \frac{1}{2}) \]

We would now like to find the matrix elements of $q$ and $p$. To do this, let us first expand $q\psi_n$ in terms of the complete orthonormal set of energy eigenfunction $\psi_n$. It can be shown
from the properties of the Hermite polynomials that this expansion contains only two terms:

\[ q \cdot \psi_n(q) = \sum_m a_{mn} \psi_m = \frac{\sqrt{h(n+1)}}{2m} \psi_{n+1} + \frac{\sqrt{hn}}{2m} \psi_{n-1} \]  

(Taking the scalar product of this equation with \( |m \rangle \), one has \((\psi_m, q \cdot \psi_n) = (|m \rangle, \sum_m a_{mn} |m \rangle) = a_{mn}\); hence the \( a_{mn} \) are indeed the matrix elements of \( q \) with respect to the basis \( \psi_n \), cf. Hageiorn Lecture Notes p.7, eq. (7).)

Multiplying (23) \( y e^{-\frac{i}{\hbar} E_n t} \) and introducing the \( \psi \)'s from (21), we get:

\[ q \cdot \psi_n = \frac{\sqrt{h(n+1)}}{2m} \psi_{n+1} e^{\frac{i}{\hbar} E_{n+1} t} e^{-\frac{i}{\hbar} E_n t} + \frac{\sqrt{hn}}{2m} \psi_{n-1} e^{\frac{i}{\hbar} E_{n-1} t} e^{-\frac{i}{\hbar} E_n t} \]

But since, according to (23),

\[ E_{n+1} - E_n = \mp \hbar \omega \]

we obtain:

\[ q \cdot \psi_n = \sqrt{\frac{h(n+1)}{2m}} \psi_{n+1} e^{i\omega t} + i\omega \frac{\hbar n}{2m} \psi_{n-1} e^{-i\omega t} \]

and similarly for \( p \)

\[ p \cdot \psi_n = \frac{\hbar}{i} \frac{\partial}{\partial q} \psi_n = \sqrt{\frac{h(n+1)m\omega}{2}} \psi_{n+1} e^{i\omega t} - i\frac{\hbar n\omega}{2m} \psi_{n-1} e^{-i\omega t} \]

Let us now define the matrix elements of an operator 0 with respect to the time-dependent functions \( \psi_n \):

\[ 0 \psi_n = \sum_m (m|0|n) \psi_m \quad (Def. \ of \ (m|0|n)) \]

Then we have:
(29) \[(n+1|q|n) = \sqrt{\frac{\hbar}{2m\nu}} \sqrt{n+1} e^{i\omega t}\]

(30) \[(n-1|q|n) = \sqrt{\frac{\hbar}{2m\nu}} \sqrt{n} e^{-i\omega t}\]

(31) \[(n+1|p|n) = \frac{i\hbar m}{2} \sqrt{n+1} e^{i\omega t}\]

(32) \[(n-1|p|n) = -i\sqrt{\frac{\hbar m}{2}} \sqrt{n} e^{-i\omega t}\]

All other matrix elements \((m|q|n)\) and \((m|p|n)\) with \(m \neq n+1\) vanish.

What does this mean? One obtains the rate of absorption and emission of electric dipole radiation by a charged harmonic oscillator placed in an electromagnetic field, if one puts these matrix elements into the classical formulas for the radiation rate. Hence a harmonic oscillator absorbs or emits radiation only with its classical frequency, and the transitions are allowed only between neighboring states. We also note that \(e^{+i\omega t}\) corresponds to an increase of \(n\) (absorption), and \(e^{-i\omega t}\) to a decrease (emission). Further, the matrix elements are practically the same for the operators \(q\) and \(p\) (in this case of the harmonic oscillator), but they are 90° out of phase (like cos and sin in the classical formulas for the variables \(q\) and \(p\)).

\[p\] and \(q\) are conjugate operators:

(33) \[pq - qp = \frac{\hbar}{i}\]

One can see this from (18) and (19), and one finds it of course also by matrix multiplication using (29) to (33). Indeed if one wants the \(q\) operator to be a simple multiplier \(q\cdot\) then by the commutation relation (3j) one has to choose \(p\) as in (18). On the other hand, one can also choose \(p\) as a multiplier, but then one is forced by (33) to use \(q = -\frac{\hbar}{\frac{1}{2} \frac{\partial}{\partial q}}\), i.e., she is using momentum space. We shall not do this here.

Now let us make the same simplification as in classical theory, introducing an operator
\( C = \sqrt{\frac{m \omega}{2}} q + i \sqrt{\frac{1}{2m \omega}} p \quad (h = 1) \)

which has a Hermitian conjugate

\( C^* = \sqrt{\frac{m \omega}{2}} q - i \sqrt{\frac{1}{2m \omega}} p \)

(since \( q \) and \( p \) are Hermitian). Evidently, \( C \neq C^* \), i.e., \( C \) is not Hermitian.

We have set \( h = 1 \). Formally, (34) is the same equation as (11), but here it applies to operators. From this definition and (26) and (27), one finds:

\( C \psi_n = \sqrt{n} \psi_{n-1} e^{-i \omega t} \)

\( C^* \psi_n = \sqrt{n+1} \psi_{n+1} e^{+i \omega t} \)

or in terms of matrix elements:

\( (n-1 | C | n) = \sqrt{n} e^{-i \omega t} \)

\( (n+1 | C^* | n) = \sqrt{n+1} e^{+i \omega t} \)

all other elements being zero. Hence \( C \) and \( C^* \) have simpler properties than \( p \) and \( q \) : \( C \) only decreases the quantum number by 1, and \( C^* \) only increases it by 1. One sees that \( C \) and \( C^* \) correspond to destruction and creation operators.

From (34), (34*) and (33) one obtains the commutation relation for \( C \) and \( C^* \) (taking care of the order of \( p \) and \( q \) in the products):

\( CC^* - C^* C = 1 \)

To obtain an expression for the energy, we find again from (34) and (34*):

\( CC^* + C^* C = \frac{p^2}{m \omega} + \omega m \frac{q^2}{2} = \frac{2}{\hbar} H \)
by comparison with (1). From this we find for $H$ (since $CC^* = 1 + C^*C$):

\[ H = \frac{1}{2} (CC^* + C^*C) = \omega (C^*C + \frac{1}{2}) \]

Now $C$ has the nice property to reduce the quantum number by 1 (and to multiply by $n$), and $C^*$, to increase it by 1 and to multiply it by $\sqrt{n+1}$, so $C^*C$ leaves the quantum number the same:

\[ C\psi_n = n\psi_{n-1} e^{-i t} \]
\[ C^*C\psi_n = n C^* \psi_{n-1} e^{-i t} = n \cdot n \psi_n = n\psi_n \]

So $C^*C$ just multiplies by $n$. If we had taken $CC^*$, we would have obtained $(n+1)$ instead, so we see again that the difference $CC^* - C^*C$ is the unit operator. Substituting for $C^*C$ in (41), we have:

\[ H = \omega (n + \frac{1}{2}) \]

The expression $H = \omega (C^*C + \frac{1}{2})$ corresponds to the classical expression $H = \omega CC^*$ (16) (in the classical case, the order of factors does not matter).

According to the fundamental ideas of quantum mechanics, the matrix elements of $p$ and $q$ are the Fourier components of the motion. For the oscillator, the result is extremely simple: there is only one frequency $(n \pm 1)q|n\rangle$ giving the $\exp(\pm i\omega t)$ Fourier components. For high quantum numbers there is little difference between $/n + 1$ and $/n$ and we find that the Fourier components of $\exp(\pm i\omega t)$ have the same sign for $q$ and the opposite sign and a factor $i$ for $p$, thus giving a $\cos \omega t$ in the former, a $\sin \omega t$ in the latter case. The variable $C$, however, has only a $\exp(+i\omega t)$ - Fourier component; $C^*$ of course has the $\exp(-i\omega t)$ - component only.
3. Quantization of a Scalar Field

To start with the simplest problem, let us first consider a scalar field (this is indeed an important case since the \( \eta \) meson field is a (pseudo-) scalar field). A classical wave field \( F(x,y,z,t) \) is specified by its amplitude at all points of space, at any time \( t \); hence one needs an infinite number of coordinates \( q \) to define the field. We do not specify anything about \( F \) except that it could in principle be measured by a test body (e.g. by observing the compression of a test proton. Since we have a scalar field, we should not use something like a force on a test body. Actually, the interaction between a meson field and a proton goes via the spin, but spin is a non-classical quantity).

To quantize this field, one has to start from the equation of motion and then to find \( L \) and \( H \). Which equation describes the classical behaviour of our field \( F \)? For the electromagnetic field, this would be Maxwell's equations, of course. Let us extrapolate from known scalar fields: if we consider the field in vacuum, without interactions, then the equation of motion should permit the propagation of waves in space, with a wave number \( k \) and a frequency \( \omega \), where:

\[
(45) \quad k^2 + \kappa^2 = \omega^2 \quad (c = 1)
\]

Here \( \kappa^2 \) is just an experimental constant (it will turn out after quantization that \( \kappa = \frac{mc}{\hbar} \), i.e., the wave can be considered as particles of Compton wave number \( \kappa \)). Now the wave equation that guarantees equation (45) and also the superposition principle, is of the form

4491/e
(46) \[ \frac{\partial^2 F}{\partial t^2} - \nabla^2 F + \kappa^2 F = 0 \] Classical equation of motion *)

This is a second order equation. Maxwell's equations on the other hand are of first order, but a second order wave equation can be derived from them by iteration. (This equation does not have the third term since light quanta have mass 0.) For all fields described by more than one component (vectors for Maxwell's equations, spinors for Dirac's equation), it is necessary to have also first order equations in order to determine the relations between the components. For the scalar field, we do not need this.

Since our equation of motion (if one neglects the term \( \nabla^2 F \)) is the same as the equation of motion of the oscillator (5), we shall expect to get oscillations again, but (due to the coupling between neighbouring points implied by the differentiation \( \nabla^2 F \)) these oscillations will propagate in space. We shall discuss below how the problem can be reduced to oscillations in a convenient manner.

At first we shall show that the equation of motion can be derived from the following Lagrangian:

(47) \[ L = \frac{1}{2} \left\{ \dot{\nabla}^2 - (\nabla F)^2 - \kappa^2 F^2 \right\} \, \text{d}r^3 \]

*) Eq. (46) is formally identical with the relativistic wave equation. One obtains this equation in a manner analogous to that used to obtain the Schrödinger equation: one substitutes the operators (18) and (19') into the relation between energy and momentum for a relativistic free particle:

(46') \[ E^2 = c^2 \vec{p}^2 + m^2 c^4 \]

and obtains:

(46'') \[ - \hbar^2 \frac{\partial^2 \psi}{\partial t^2} = - \hbar^2 c^2 \nabla^2 \psi + m^2 c^4 \psi \]

This equation has solutions of the form \( e^{i(\vec{k} \cdot \vec{r} - \omega t)} \) if \( k^2 + \left( \frac{mc}{\hbar} \right)^2 = \left( \frac{\omega}{c} \right)^2 \) as one can verify by substitution.
i.e. we have to show that:

\[ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \]

is the equation of motion. Now our coordinate \( q \) is the value of \( F \) at each point, hence \( \frac{\partial L}{\partial q} \) (which by definition is the conjugate momentum \( \Pi \)) is given by:

\[ \Pi = \frac{\partial L}{\partial F(x,y,z)} \quad \text{(at a certain point)} \]

Considering the integral simply as a sum over the values at all points, we have:

\[ \frac{\partial L}{\partial F} = \dot{F}(x,y,z) \]

or:

\[ \dot{F}(x,y,z,t) \leftrightarrow F(x,y,z,t) \]

where the symbol \( \leftrightarrow \) means that they are canonically conjugate.

Comparing this to the conjugate momentum (3) in the case of the harmonic oscillator, we see that we have an oscillator of symbolic mass \( m = 1 \). The dimensions, of course, are different, but from the very beginning, \( F \) has not had the dimension of a length.

To obtain \( \frac{\partial L}{\partial q} \), we have to resort to a partial differentiation:

\[ \int (\nabla F)^2 \, dr^3 = \int \left( \frac{\partial F}{\partial x} \frac{\partial F}{\partial x} \right) \, dx \, dy \, dz + \ldots \]

where:

\[ \int \frac{\partial F}{\partial x} \cdot \frac{\partial F}{\partial x} \, dx = F \left. \frac{\partial F}{\partial x} \right|_{-\infty}^{+\infty} - \int F \frac{\partial^2 F}{\partial x^2} \, dx, \text{ etc.} \]

so that:

\[ \int (\nabla F)^2 \, dr^3 = -\int F (\nabla^2 F) \, dr^3. \]

Differentiating this term with respect to \( F \), we get the same thing twice (once from each factor \( \nabla F \)).
\[ (54) \quad \frac{\partial}{\partial F} \int (\nabla F)^2 \, dr^3 = -2 \nabla^2 F \]

(In somewhat greater detail: consider a variation of \( F \) in the vicinity of \( r \).)

Then:
\[ \frac{\partial}{\partial F} \int (\nabla F)^2 \, dr^3 = \frac{\int (\nabla F + \delta F)^2 \, dr^3 - \int (\nabla F)^2 \, dr^3}{\int \delta F \, dr^3} = \frac{\int 2 \nabla F \cdot \nabla (\delta F) \, dr^3}{\int \delta F \, dr^3} = \]

(and by partial integration)
\[ = -2 \frac{\int \delta F \cdot \nabla^2 F \, dr^3}{\int \delta F \, dr^3} = -2 \nabla^2 F \]

since \( \nabla^2 F \) is constant over the small range in which \( \delta F \neq 0 \).)

Hence:
\[ (55) \quad \frac{\partial L}{\partial F} = -\nabla^2 F - \kappa^2 F \]

Introducing (49) and (55) into (48), we obtain indeed the wave equation (46). Hence the Lagrangian (47) is correct.

To derive the Hamiltonian, we need the general rule:
\[ (56) \quad H = \sum_i p_i \dot{q}_i - L \quad (= 2T - L) \]

where the sum now means integration over all points of space. We obtain using (49) and (47):
\[ (57) \quad H = \int \dot{r}^2 \, dr^3 - L = \frac{1}{2} \left\{ \frac{1}{2} \int \dot{F}^2 + (\nabla F)^2 + \kappa^2 F^2 \right\} \, dr^3 \]

It is not yet obvious that this is an oscillator Hamiltonian. To be sure, the first term corresponds to the kinetic, the last to the potential energy of the oscillator (1), but there is also the middle term which takes into account differences of \( F \) at neighbouring points, hence the oscillators are coupled. But as in all cases of coupled classical oscillators, it is possible to introduce normal vibrations that vibrate in an uncoupled manner (as in the case of the vibrating string, where the points of the string are coupled to each other, but one introduces the harmonics which are not). Now this can be done in many ways, since there are many degeneracies: e.g., for the normal vibration with frequency \( \omega \), one has the choice between the
running and standing plane waves in all directions, spherical waves, etc. The best choice depends on the problem. Since Dr. Pupini talks about the fixed source meson problem, we shall first use spherical waves as our normal vibrations.

We expand the function $F$ in terms of some complete orthogonalized system (in the next lecture, the meaning of this expansion will be considered in greater detail):

\[ F(xyz) = \sum_i q_i(t) f_i(xyz) \]

where the $q_i$ should be our normal coordinates and where:

\[ \int f_i f_j^* \, d^3r = \delta_{ij} \]

We choose $q_i$ and $f_i$ real and in such a way that every term $q_i f_i$ of eq. (48) obeys the equation of motion (46) \(^*\), by substituting $q_i f_i$ into (46):

\[ \dot{q}_i f_i - q_i \nabla^2 f_i + \kappa^2 q_i f_i = 0 \]

\[ \dot{q}_i \equiv \frac{\nabla^2 f_i}{f_i} - \kappa^2 = \omega_i^2 \]

where we have called the arbitrary separation constant $-\omega_i^2$. Then:

\[ q_i = a_i \cos \omega_i t + b_i \sin \omega_i t \]

and $f_i$ must obey the Poisson equation:

\[ \nabla^2 f_i = - (\omega_i^2 - \kappa^2) f_i = - k_i^2 f_i \]

(\(\text{using (45)}\)) where $k_i$ is still an arbitrary constant.

\(^*\) Though $q_i f_i$ is an eigenfunction of (46) in the mathematical sense, obtained by imposing a boundary condition (65) upon a differential equation, it is not a state function of the system in the quantum-mechanical sense. Indeed we are still dealing entirely with classical quantities.
Example 1

We choose spherical waves for our functions $f$. As a mathematical convenience, we enclose our system in a large sphere of radius $R$. We use spherical coordinates. Then if we require $f(R)=0$, we obtain solutions of the Poisson equation only for discrete values of $k = k_i$; and there are several solutions for every $k_i$ distinguished by two constants $l$ and $m$ (which, in particle mechanics, would correspond to the angular momentum and its $z$-component):

$$f_{n\ell m} = \frac{1}{r} J_{\ell} (k_n r) Y_{\ell m} (\theta, \varphi)$$

with

$$f(R) = 0$$

where the $J_{\ell}$ are Bessel functions, and the $Y_{\ell m}$ are spherical harmonics. For $r \to \infty$,

$$J_{\ell} (k_n r) \to \left( \frac{2}{k_n} \right)^{\frac{\ell + 1}{2}} \sin (k_n r - \frac{\pi \ell}{2}).$$

To satisfy (65), we must have approximately:

$$k_n r - \frac{\pi \ell}{2} = n \pi \quad (n = 1, 2, \ldots)$$

or:

$$k_n = \frac{(n + \frac{\ell}{2}) \pi}{R}$$

which gives us the discrete set of wave numbers. Actually, $k$ would also depend on $l$; but since in practice one will only use wavelengths $\ll R$ in a laboratory of dimension $R$, one will have $kR \gg 1$, whereas $\ell$ will usually be small (0 or 1, 2, etc.). Hence $k$ depends essentially on $n$.

Now let us introduce the expansion (56) with $f_1 = f_{n\ell m}$ into the Hamiltonian (57):

$$H = \frac{\hbar}{\kappa} \int \left\{ \left( \sum q_1 f_1 \right)^2 + \left( \nabla q_1 \nabla f_1 \right)^2 + \kappa^2 \left( \sum q_1 f_1 \nabla f_1 \right)^2 \right\} \varphi^2 =$$
\[
\sum_{ik} \left[ q_i q_k \int f_i f_k \, dr^3 + q_i q_k \int (\nabla f_i)(\nabla f_k) \, dr^3 + \kappa^2 q_i q_k \int f_i f_k \, dr^3 \right] = J_{ik} = d_{ik}
\]

But now:

\[(70) \quad \int (\nabla f_i)(\nabla f_k) \, dr^3 = -\int (\nabla^2 f_i) f_k \, dr^3 = + \kappa_n^2 \int f_{ik} \]

by partial integration, and using (63). So we obtain:

\[(71) \quad H = \frac{1}{2} \sum_i \left[ q_i^2 + \kappa_n q_i^2 + \kappa^2 q_i^2 \right] = \frac{1}{2} \sum_i \left[ \omega_{i}^2 + \omega_{i} q_i^2 \right] = \frac{1}{2} \sum_i (\omega_{i}^2 + \omega_{i}^2 q_i^2) \]

since:

\[(72) \quad \omega_{i} = \kappa_n^2 + \kappa^2 \]

according to (45). Here, i stands for the triple index (n l m).

We see from this equation that we have reduced the Hamiltonian H from form (63) which contained the coupling between oscillations at neighbouring points, to a new form which is the sum over the Hamiltonians of independent oscillators of mass 1, each one described by a normal coordinate q_i. For l = 0, e.g., we have Y = const; f is a standing wave of shape \(\sin \frac{kr}{r}\), etc.

Now the quantization of the field can be performed in a single step. Since we know the eigenvalues of a quantized oscillator (22), we obtain from (71) for the energy of the field:

\[(73) \quad H = \sum_i \hbar \omega_{i} \left( N_i + \frac{1}{2} \right) \]

where \(N_i\) is the number of quanta in the oscillator of frequency \(\omega_{i}\) (occupation number). The first remarkable point about this equation is that even in the lowest state where all \(N_i = 0\), we have an infinite zero point energy. This fact is not surprising and always occurs with fields because of the infinite number of degrees of freedom. It has nothing to do with the divergent infinities that come from the coupling and it presents no difficulty since the value is constant and might even be subtracted out.
There is, however, also a zero point amplitude since our oscillator can never be completely quiet. That means that the so-called classical vacuum (where $F = 0$ everywhere) does not exist. Or rather, there are two possibilities for a vacuum: one is to make $F = 0$ everywhere for one extremely short instant (which is possible), but then because of the uncertainty principle since the $q_i$ are completely determined, the $\dot{q}_i$ are completely undetermined, hence all components appear in the oscillation immediately afterwards, and $H$ (71) diverges. Hence to define a vacuum one ordinarily uses the lowest state of energy, which is a stationary state as opposed to the first definition. But in such a vacuum, the expectation value of $F^2$ at any point is again infinite because each of the infinitely many oscillators adds a finite positive amount.

Here there is a real problem. The vacuum now is not empty. Applying this idea for a moment to the electromagnetic field and putting an electron into it, e.g., the electron will be shaken around by the vacuum oscillations. For a point electron these oscillations would even diverge. Hence the attempts at renormalization. For an extended electron, there will be no effect of the high frequencies since they will average out over the size of the electron. There remains, however, a finite effect.

Can one measure the zero point energy, e.g. of an oscillator? Yes, if the oscillator can be taken apart (like the $H$ atom, by ionization). But the field oscillators cannot be taken apart, hence one cannot make a real vacuum.

But even if the zero point energy is not observable, the phenomena connected with it are observable, e.g. by their influence on an electron. The best example of this is the Lamb shift: one can show in a simple way *) that the oscillations of an electron are different if an external electric field is superimposed upon the vacuum. The energy difference between the oscillations in this case and those in vacuum is finite. For the Coulomb field, this difference is the Lamb shift.

Third Lecture

(Notes collected by T. Fazzini and H. Paul)

February 13th, 1958

We shall start with an illustration of what has been done the last time, to make more plausible why \( x, y, z \), which are normally coordinates, are not considered coordinates for our present purpose.

To do this, we assume that space is made up of discrete points \( \ldots \ldots \) (a sort of crystal space). Then \( F \) is defined only at these discrete points: \( F(x_i) \). Our problem is similar to that of the quantization of motion of a crystal. In this case, the variable would be the displacement at a certain point \( i \); in our case, it is more generally the value of a scalar function at a point \( i \):

\[
F(x_i) = F_i
\]

where space appears only as an index. The Hamiltonian (57) would then become a sum over the different space points:

\[
H = \sum_{i} \left\{ \Pi_i^2 + (\nabla F)_i^2 + \mu^2 F_i^2 \right\}
\]

(\( \Pi \) stands generally for the canonical conjugate of the coordinate. In our case, \( \Pi = F \)).

As before, \( \nabla F \) represents a coupling between the field values at neighbouring points, and stands for \( (\nabla F)_i = (F_{i+1} - F_i)/\delta \), where \( \delta \) is the distance between points. When solving the problem of crystal vibrations one introduces normal coordinates:

\[
q_k = \sum_{i} F_i a_k^i
\]
which are suitably chosen to give independent vibrations. Evidently, \( q_k \) is just a linear combination of the values of \( F \) at all points. For a continuous space, this goes over into the integral

\[
q_k = \int F(x) a_k(x) \, dx^3
\]

where the index \( i \) is now replaced by the continuous coordinate \( x \), so that the transformation function has only the index \( k \) of the normal coordinate (this index was called \( i \) in the last lecture). The \( a_k(x) \) form a matrix with one discrete and one continuous index.

To make the coordinates \( q_k \) independent (as \( F_i \) are), the transformation matrix \( a_k^i \) in (76) must be unitary (here, it is even orthogonal, since we use real coordinates). Or, for continuous space, the transformation functions \( a_k(x) \) (called \( f_k \) last time) must be orthonormal. The suitable choices for \( a_k(x) \), which give rise to normal coordinates are the solutions of the wave equation. For example:

\[
a_k(x) = f_k(x) = \sqrt{\frac{2}{R}} \frac{1}{r} J_\ell(k_n r) Y_{\ell m}
\]

where \( k \) stands for \((n\ell m)\). So we have instead of (77):

\[
q_k = \int F(x) f_k(x) \, dx^3
\]

Remark: we assume the \( f_k \) to be real, though the \( Y_{\ell m} \) are actually complex. But it is very easy to make them real: since the complex terms which they contain are only of the form \( e^{im\varphi} \) and \( e^{-im\varphi} \), it suffices to take terms of equal \( m \) (but opposite sign) together and combine them to \( \cos m\varphi \) and \( \sin m\varphi \). This means that we use standing waves rather than travelling waves.

Now because of the orthogonality (*) of the \( a_k^i \), (76) can

*) The relations between the functions \( a_k(x) \) which correspond to the orthogonality of the matrix \( a_k^i \), are the following:

\[
\sum_i a_k^i a_k^i = \delta_k^k
\]

\[
\int a_k(x) a_{k'}(x) \, dx = \delta_k^{k'}
\]

\[
\sum_k a_k^i a_k^j = \delta_{ij}
\]

\[
\int_k a_k(x) a_{k'}(x') \, dx = \delta(x-x')
\]

where the first is the orthonormality relation, and the second guarantees that the system is complete.
be solved for $F_i$ to give

\[(80) \quad F_i = \sum_k q_k a_k^i\]

Similarly, (79) can be solved for $F(x)$ (indeed, the $q_k$ are just the Fourier components of $f(x)$):

\[(81) \quad F(x) = \sum_k q_k f_k(x).\]

This is the Fourier expansion of $F$ (which we had as (58) last time). The amplitude $q_k$ are our new coordinates, and the $f_k$ the transformation parameters (which transform from a continuous to a discrete system); the $x$ are now just indices, not coordinates.

With this expansion for $F(x)$, we can again rewrite the classical Hamiltonian as in (71):

\[(82) \quad H = \sum_n (p_k^2 + \omega_k^2 q_k^2)\]

where $p_k$ is the canonically conjugate of $q_k$. (Here, $p_k = \dot{q}_k$; to show this, substitute the expansion (81) into the expression for $L$, (47), thereby replacing the index $xyz$ by the index $k$, and use $p_k = \frac{\delta L}{\delta \dot{q}_k}$.)

We now proceed to the quantization of the Hamiltonian, in the pure Schrödinger sense, replacing the $q_k$ and $p_k$ in $H\Psi = E\Psi$ by their operators. The state function $\Psi$ now will be a function of the coordinates $q_1$, $q_2$, ... Because of the form of the Hamiltonian, we get the Schrödinger equation for independent oscillators; it has solutions which are simple products:

\[(83) \quad \Psi (q_1, q_2, ...) = h_{n_1}(q_1) h_{n_2}(q_2) \ldots\]

where the Hermite polynomials $h$ are the ordinary oscillator eigenfunctions. The numbers $n_1$, $n_2$, ... are the excitation numbers of the different oscillators. For the corresponding energy we again have the sum (73):

\[(84) \quad E = \sum_k (n_k + \frac{1}{2}) \hbar \omega_k\]
Since it is disagreeable to write infinite products for \( \Psi \) (especially since usually many excitation numbers are zero), we introduce a shorthand for (83):

\[
\Psi \equiv \Psi(n_1, n_2, \ldots).
\]

The commutation relation of the Fields

Now we wonder what the various operators do to our wave function, i.e., we are interested in the matrix elements of these operators. Since \( q_i \), e.g., only operates on the \( i \)th oscillator, one obtains directly from (26):

\[
q_i \Psi(n_1, n_2, \ldots) = \sqrt{\Psi} (\ldots n_i + 1 \ldots) e^{i\omega_1 t} + \\
\sqrt{\Psi} (\ldots n_i - 1 \ldots) e^{-i\omega_1 t}
\]

If one replaces the coordinates \( q_k \) in (81) by the operators \( q_k' \), then the field \( F(x) \) also becomes an operator (with index \( x \)). (The transformation functions \( f_k(x) \) are just ordinary numbers, "c-numbers" so they remain unchanged.) So the product

\[
F(x) \Psi(n_1 \ldots) = \sum q_k f_k(x) \Psi(n_1 \ldots) = \ldots
\]

has a well-defined meaning (one would have to substitute the \( q_k \Psi \) from (86), of course, to get the complete expression).

Similarly for the operator

\[
\Pi(x) = F(x) = \sum_k q_k f_k(x) = \sum_k p_k f_k(x)
\]

we can find

\[
\Pi(x) \Psi(n_1 \ldots) = \sum_k p_k f_k(x) \Psi(n_1 \ldots) = \ldots
\]

by using the relation (27).
Now what is the commutation relation \([\Pi(x), F(x')] = \Pi F - F\Pi\)? This is now well-defined since the fields \(F\) and \(\Pi\) are operators that operate on the state function. Using the expansion (81) and (88) and the commutation relation (33), we obtain:

\[
(90) \quad [\Pi(x), F(x')] = [\Sigma p_k f_k(x), \Sigma q_i f_i(x')] = \\
= \sum_i \left[p_i q_i (x)f_i(x') \right] \\
= \frac{\hbar}{i}
\]

since all other terms drop out. Next we use the completeness relation (cf. the footnote on p. 17)

\[
(91) \quad \sum_i f_i(x) f_i(x') = \delta(x-x')
\]

which holds for every complete system, and we obtain the relation

\[
(92) \quad [\Pi(x), F(x')] = \frac{\hbar}{i} \delta(x-x')
\]

This is the relation from which theoreticians usually start out to quantize a field. Working backwards, one can indeed find state functions and energy eigenvalues from it. We note that \(F(x)\) commutes with \(F(x')\), and \(\Pi(x)\) with \(\Pi(x')\) since they are linear combinations of commuting operators, i.e., one can measure the field \(F(x)\) in all space at one time, or one can measure the field \(\Pi(x) = \hat{F}(x)\) in all space at one time; one can also measure the two fields at different points at the same time, but one cannot measure both fields at the same point at the same time.

It should be obvious that the functions \(f_i\) are just a transformation of the coordinates and have nothing to do with the eigenfunction (33) of the system.

As in the case of the simple harmonic oscillator, it is again useful to express \(F(x)\) by \(C_k\) and \(C^*_k\) (defined exactly as in (34)), instead of only by the \(q_k\). We use (14):

\[
(93) \quad q = \frac{1}{\sqrt{2\omega}} (C + C^*)
\]
and obtain:

$$F(x) = \sum_k \frac{1}{\sqrt{2\omega_k}} (C_k + C_{-k}^*) f_k(x).$$

\[\text{creation part}\]
\[\text{annihilation part}\]

**Remark**: We have chosen real $f$'s. As we know, the choice is not unique, since $\omega_k$ depends only on $n$, but not on $l$ and $m$. If we therefore choose complex $f$'s, i.e., the ordinary $Y$'s, or running plane waves as is customary, we shall obtain a slightly different equation for $F(x)$.

$$F(x) = \sum_k \frac{1}{\sqrt{2\omega_k}} (C_k f_k + C_{-k}^* f_{-k}^*).$$

If we consider again spherical waves for $f$ as in (78) but with the ordinary complex meaning of $Y_{lm}$, we note that many functions appear twice since $f_{nlm}^* = f_{nl,-m}$ so that two terms might seem to mean the same. This is not so, as can be shown by considering the time-dependance of $C$ and $C^*$ (cf. (35) and (36)). Let us define $k \equiv (nlm), -k \equiv (nl,-m)$ and let us write down the parts with positive and negative $m$ separately:

$$F(x) = \sum_{k>0} \frac{1}{\sqrt{2\omega_k}} \left[ (C_k f_k + C_{-k}^* f_{-k}^*) + (C_{-k} f_k + C_k^* f_{-k}^*) \right]$$

where the $f$'s contain: $e^{+i\varphi}, e^{-i\varphi}, e^{-i\varphi}, e^{+i\varphi}$ and the $C$'s produce: $e^{-i\omega t}, e^{-i\omega t}, e^{-i\omega t}, e^{+i\omega t}$

hence we obtain:

$$e^{i(m\varphi-\omega t)}, e^{-i(m\varphi-\omega t)}, e^{-i(m\varphi+\omega t)}, e^{+i(m\varphi+\omega t)}$$

which means: rotation in the positive sense

rotation in the negative sense

Recalling that the $f_k$ have been chosen with real $r$- and $\theta$-parts, we note that the first two terms refer to spherical waves standing in the $r$- and $\theta$-directions, and rotating in the positive sense about the $z$-axis, and similarly the last two terms refer to rotation in the negative sense about the $z$-axis. Hence the first half of the sum contains the complex conjugate parts of positive rotation, and the second half of negative rotation.
four terms as annihilation of a quantum with positive rotation, creation of a quantum with positive rotation, annihilation of a quantum with negative rotation, and creation of a quantum with negative rotation.

We will derive equation (95) for the case of the complex $Y_{\ell m}$. We start with the real $f_k$ and have then, instead of the pair $f_{n\ell m}, f_{n\ell - m}$ the real pair $f_{n\ell |m}^c = \frac{1}{2} (f_{n\ell m} + f_{n\ell - m}) \equiv f^c_{n\ell |m}$

$$f_{n\ell |m}^s = \frac{1}{2i} (f_{n\ell m} - f_{n\ell - m}) \equiv f^s_{n\ell |m}$$

Let us take out of the sum (95) just the part for one pair $n\ell |m|:

$$q_c f^c + q_s f^s$$

This can be written in the form:

$$(96) \quad q_c f^c + q_s f^s = (q_c - iq_s) f_{n\ell + m} + (q_c + iq_s) f_{n\ell - m}$$

We decompose each $q$ in the usual form (cp (14)):

$$(97) \quad q_c = \frac{1}{\sqrt{2\omega}} (c_c + c^*_c), q_s = \frac{1}{\sqrt{2\omega}} (c_s + c^*_s)$$

We now introduce now $C$'s by:

$$(98) \quad \frac{1}{\sqrt{2}} (c_c - ic_s) = C_+, \quad \frac{1}{\sqrt{2}} (c^*_c + ic^*_s) = C^*_+$$

$$\frac{1}{\sqrt{2}} (c_c + ic_s) = C_-, \quad \frac{1}{\sqrt{2}} (c^*_c - ic^*_s) = C^*_-$$

The new $C_\pm, C^*_\pm$ have the same characteristic properties as the $c_c, c^*_c, c_s, c^*_s$, as to time dependence and commutation relations. By putting (98) into (97) and (97) into (96) we get $q_c f_c + q_s f_s = c_+ f_+ + C^*_+ f^*_+ + C_- f_- + C^*_- f^*_-$ where $f_\pm = f_{n\ell \pm m}$.

Observing that $f^*_\pm = f^*_+$, we get:

$$q_c f_c + q_s f_s = c_+ f_+ + C^*_+ f^*_+ + C_- f_- + C^*_- f^*_-$$

which leads to (95)!
Momentum and Angular Momentum of the Field

We want now to define the linear momentum of the field. This is not Π(x), the canonical conjugate of F(x), which is a generalized momentum. Rather we want to define what a classical physicist would understand by the momentum of the field.

Hence we really want the conjugate of x:

\[ P = \frac{\hbar}{i} \frac{\partial}{\partial x} \].

The meaning of this is not immediately obvious since P should act on ψ, but x is contained in ψ only as an index, not as a coordinate. Indeed, ψ is a function of the q's:

\[ ψ = ψ(q_1, q_2, \ldots) \]

But, (as we see more simply from the case of a discrete space) the q's are linear combinations of the F's at the various points (76), hence we can also write ψ as a function of the F_j = F(x_j):

\[ ψ = ψ(F_1, F_2, \ldots) \]

So ψ is really a function of the fields at each point of the space.

We now apply the displacement operator \( P = \frac{\hbar}{i} \frac{\partial}{\partial x} \) to this object. Since it is the F's that depend on x, and ψ depends on the F's we have the situation similar to that of the differentiation of an implicit function: first we have to differentiate the state function with respect to all the F_j, and then to differentiate every F_j with respect to x, (which means taking the differences at two neighbouring space points) and to divide by Δx:

\[ P = \frac{\hbar}{i} \sum_j \frac{F_{j+1} - F_j}{\Delta x} \frac{\partial}{\partial F_j} \]  

(99)

Returning to continuous space, the difference quotient becomes \( \frac{\partial F}{\partial x} \) at point x, the expression \( \frac{\hbar}{i} \frac{\partial}{\partial F_j} \) is by definition *) equal to Π at point x, and instead of summing, we have to integrate over all points of space:

\[ \int \frac{\partial}{\partial F_j} \]

*) Since Π and F are defined as canonically conjugate operators, we have quite generally \( Π = \frac{\hbar}{i} \frac{\partial}{\partial F} \). In this special case of the oscillator, we have in addition \( Π = \frac{\partial}{\partial F} \).

5665/e
This is the definition of the momentum operator, which is meaningful since we know $F(x)$ and $\Pi(x)$.

Using this operator we shall now calculate the momentum contained in the field. Clearly the spherical functions $f_k$ (78) would be inconvenient for this purpose, so we shall use running plane waves instead (example 2).

Using $f_k = e^{\frac{ik_x}{2}}$ (for simplicity, we assume a unit normalization volume) we obtain from (95):

$$F = \frac{1}{\sqrt{2\omega_k}} \left( C_k e^{\frac{ik_x}{2}} + C^*_k e^{-\frac{ik_x}{2}} \right)$$

We need its derivative with respect to $x$:

$$\frac{\partial F}{\partial x} = i \sum_k \frac{k_{kx}}{\sqrt{2\omega_k}} \left( C_k e^{\frac{ik_x}{2}} - C^*_k e^{-\frac{ik_x}{2}} \right)$$

and also the derivative with respect to time (using the time dependence (35, 36) of $C$ and $C^*$):

$$\Pi = F = i \sum_i \frac{\omega_i}{\sqrt{2}} \left( - C_i e^{\frac{ik_x}{2}} + C^*_i e^{-\frac{ik_x}{2}} \right)$$

Introducing (102) and (103) into (100), all products of exponentials drop out after integration because of the orthonormality of the $f_k$ except those products where one factor is exactly the complex conjugate of the other:

$$P_x = \frac{i}{\hbar} \sum_k k_{kx} \left( C_k C^*_k + C^*_k C_k \right) = \sum_k k_{kx} \left( C^*_k C_k + \frac{1}{\hbar} \right) (\hbar=1)$$

where we have used the commutation relation (39). But from (43) we know that the eigenvalue of $C^*C$ is $n$. Also, since there are moments in all directions (to every $k$ there is another $k$ of opposite sign), it can be made somewhat plausible that the $\frac{1}{\hbar}$'s in the sum will cancel out (independent of whether or not there is a photon in a particular state). So we have:

$$P_x = \sum_k k_{kx} n_k \quad (\hbar = 1)$$
and this is indeed the result we would have expected.

The calculation for the angular momentum is entirely similar. We use, of course, again the spherical $f_k$'s and obtain:

$$L_z = \int \frac{\partial F}{\partial \phi} \Pi \, d\mathbf{r}$$

(106)

Since we have now a differentiation with respect to $\phi$, and since the $f_k$'s contain the spherical harmonics, we get $m_k$ where we got $k_k$ in the preceding calculation. So the result is:

$$L_z = \sum_k m_k n_k$$

(107)

again as expected.
To see what is really the meaning of a field one has to consider its interaction with some matter. For the sake of simplicity one can start with the interaction of a scalar field with one particle, describing the motion of this particle by classical mechanics. If we start treating classically the field also, we can write the total Lagrangian adding an interaction term \( J \) to the sum of the Lagrangian \( L_0 \) of the free field (eq. 47) and the Lagrangian \( L \) part of the particle moving perhaps in an external field of force.

\[
L = L_0 (F) + L \quad (q, \dot{q}, F(x), \dot{F}(x)) \quad (10^e)
\]

Here \( F(x) \) is the field amplitude at the point \( x \) (remember that \( x \) has to be considered as a parameter or an index rather than a dynamical variable) and \( q \) is the coordinate of the particle.

The interaction term \( J \) has to depend on the dynamical coordinates \( q \) and \( F(x) \) of the interacting quantities and could depend on their time derivatives \( \dot{q} \) and \( \dot{F}(x) \), but we will restrict ourselves to the simplest case in which it does not; so we can interpret \( J \) as the interaction energy between particle and field as can be seen from

\[
H = \sum_k \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} - L = H_0 + H + J (q, F(x)). \quad (109)
\]

There are restrictions on the form of \( J \) due to physical principles, so it has to be a scalar to be relativistically invariant, (in particular, invariant under spatial rotations) and it has to depend only on the value of \( F(x) \) at the point \( q \) where the particle is, to satisfy locality.

The simplest choice is the "scalar coupling"

\[
J = g F \quad (x=q)
\]

which is better written in the form

\[
J = g \int F(x) \int (x-q) \, d^3 \quad (110)
\]

\( g \) is of course the coupling constant.

One can see what is the effect of this interaction term on the equation of motion of the particle by inserting (109) in the equation

\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = \gamma \quad (111)
\]

One gets from the \( L \) part. term

\[
m \ddot{q} = F \text{ ext.}
\]
where $F_{\text{ext.}}$ is the external force; and the coupling term adds to the right a term

$$-\frac{\partial J}{\partial q} = -g \int_{F(x)} \frac{\partial}{\partial q} \delta(x-q) \, dx^3 = g \int_{F(x)} \frac{\partial}{\partial x} \delta(x-q) \, dx^3 =$$

$$-g \left. \left( \frac{\partial F(x)}{\partial x} \delta(x-q) \right) \, d^3x = -g \left. \frac{\partial F(x)}{\partial x} \right|_{x=q};$$

so the equation of motion of the particle becomes

$$m \ddot{q} = F_{\text{ext}} - g \left. \frac{\partial F(x)}{\partial x} \right|_{x=q} \quad (112)$$

This equation shows that $F$ acts on the particle through its gradient like e.g. an electric potential on a charged particle (but one has to remember that in a relativistic treatment $F(x)$ is a scalar whereas the electric potential is the fourth component of a 4-vector).

As one can give a physical significance to a quantity only by assigning a procedure to measure it, one can use (112) to give a meaning to the field $F$ as it allows its gradient to be measured by the acceleration of the particle.

In the same way as for the particle one can look at the effect of the interaction on the equation of motion of the field and one gets (compare page 10-11 of 2nd lecture)

$$(\nabla^2 - k^2) \Phi - \Phi = \frac{\partial J}{\partial F} = g \delta(x-q) \quad (113)$$

where the right hand side represents an inhomogeneity in the equation, that is, the particle not only suffers the action of the field, but also acts as a source producing the field, and for a moving particle one finds here the simplest case of radiation.

A static solution of equation (113) is found by taking its Fourier transform (after having put $\Phi = 0$ and $q = 0$)

$$(-k^2 - \lambda^2) \hat{F}(\lambda) = \frac{1}{4\pi} \delta$$

that is

$$F(\lambda) = \frac{-\delta}{4\pi} \frac{1}{\lambda} \quad (114)$$

And this corresponds to an $F(x)$ given by

$$F(x) = \frac{-\delta}{4\pi} \frac{e^{-kx}}{r} \quad \text{(Yukawa solution)} \quad (115)$$

Such a calculation is possible only for a classical particle where one can put $q = 0$ all the time, but not for the quantum-mechanical case, where putting $q = 0$ at a certain time implies having very large momentum present so that one can no longer speak of a static solution.
One can still have this static solution only in the limiting case of a very heavy particle for which large momentum still does not mean high velocity.

To see what happens in this limit when one tries to quantize the field, one can take again the Fourier transform of (115) but without putting $F = 0$

\[ -\omega^2 F'(k) - \omega^2 F(k) = \xi; \]

or

\[ -\omega^2 (F(k) + \frac{\xi}{\omega^2}) - \omega^2 F(k) = 0; \quad (116) \]

Classical mechanics tells us that this is the equation of motion of an oscillator with the centre displaced at the point $-\frac{\xi}{\omega^2}$ and as there is no dynamical variable involved other than $F(k)$, the quantization is easily done in the same way as for the free field. But the result is still a static solution superimposed on the running waves of the free field, and this means that these running waves are not scattered, that is, there is no scattering of the quanta of the scalar field from a very heavy particle; this result is just due to the fact that no dynamical variable of the particle is present in the equation of motion of the field. An analogous result is obtained in calculating Compton scattering from a charged particle with $M \to \infty$.

If one uses the Yukawa solution (115) due to a particle at rest at $q=0$ as the field to insert in the equation of motion (112) of a particle at $q=r$ one gets (in the same approximation of infinite mass) the force exerted by the first particle on the second one. This force can be expressed by a central purely attractive potential:

\[ V(r) = -\frac{\xi^2}{4\pi} \frac{e}{r} \quad (117) \]

One can get this result also by inserting the field (115) into the interaction energy (110) of the second particle. This interaction is not a good one to be used for describing meson interactions because as we have seen it gives too simple an interaction between nucleons, and, what is more important, it gives no scattering between mesons and nucleons.

The next type of coupling one can study is the "pseudovector coupling."

\[ -\lambda \mathbf{\sigma} \cdot \mathbf{\sigma} = \mathbf{\sigma} \cdot \mathbf{\sigma} \delta(x) \delta(x-q) \, dx^3 \quad (118) \]

where $\mathbf{\sigma}$ is the spin of the particle, $\delta(x)$ is a pseudoscalar field, $\lambda$ is the mass of the meson ($\mu = \lambda$).
\[ \vec{J} \] is an angular momentum i.e. an axial vector, and is a polar vector, their scalar product gives a pseudoscalar; but \( J \) being a term of a Lagrangian has to be a scalar so \( \phi \) has also to be a pseudoscalar. We see here once again that the nature of the field comes out only through its interactions.

The presence of the mass in the denominator is just due to dimensional reasons, to compensate for the extra length\(^{-1}\) introduced by the gradient.

It is important to notice that in this interaction term there are two dynamical variables of the particle, namely its position \( q \) and its spin \( \vec{\sigma} \).

A rather familiar example of such a pseudo-vector coupling is the interaction energy of a magnetic dipole \( \vec{\mu} \) in a magnetic field \( \vec{H} \) described by its magnetic potential \( G \)

\[
E = - (\vec{\mu} \cdot \vec{H}) = (\vec{\mu} \cdot \nabla) G \quad (119)
\]
\[
\vec{H} = - \nabla G \quad (120)
\]

As \( \vec{H} \) is an axial vector \( G \) is a pseudo-scalar and so we get a complete correspondence between (118) and (119) namely \( G \) corresponding to field \( \phi \) and \( \rho \) to \( \frac{\vec{\mu}}{\mu} \vec{\sigma} \).

We can use this correspondence, to get the effect of the field on the particle, by "translating" the equation of motion of the dipole

\[
\frac{\partial \vec{\mu}}{\partial t} = \frac{\vec{e}}{\mu} (\vec{\mu} \times \vec{H}) \quad (121)
\]
\[
\frac{\partial \vec{\sigma}}{\partial t} = \frac{\vec{e}}{\mu} (\vec{\sigma} \times \nabla) \phi \quad (122)
\]

Now we look at the effect of the particle on the field; this is done in the usual way by adding to the equation of motion of the free field a right hand side given by \( \frac{\partial \vec{J}}{\partial \phi(x)} \) and this turns out after an integration by parts

\[
\frac{\partial \vec{J}}{\partial \phi(x)} = - \frac{\vec{e}}{\mu} (\vec{\sigma} \cdot \nabla) \vec{\phi} (x-q) \quad (123)
\]

so we obtain

\[
(\nabla^2 - \vec{\kappa}^2) \phi - \vec{J} = - \frac{\vec{e}}{\mu} (\vec{\sigma} \cdot \nabla) \vec{\phi} (x-q) \quad (124)
\]

If we look for a static solution of this equation we have a difficulty coming from the fact that there are two dynamical variables \( \vec{\sigma} \) and \( \vec{\phi} \) of the particle involved and even if we use the limit \( M \rightarrow \infty \) that allows us to consider \( q \) as a constant we are left with \( \vec{\sigma} \).

Only one component of \( \vec{\sigma} \) can be constant, but then the other two will oscillate and we can only know their time average which is zero;
so all we can do is to calculate a time average of $\phi$. We do this again by putting $\dot{\phi} = 0$, $q = 0$ and by taking the Fourier transform of what is left of equation (124):

$$-\omega^2 \phi(k) = \frac{ie}{4\mu} (\mathbf{\sigma} \cdot \mathbf{k})$$  \hspace{1cm} (125)

Dividing by $-\omega^2$ and making the inverse transformation we get finally

$$\phi(r) = -\frac{e}{4\pi \mu} (\mathbf{\sigma} \cdot \nabla) \frac{e^{-kr}}{r}$$  \hspace{1cm} (126)

We now calculate the forces that will result from such a model between two nucleons (the result will not be exact, because the spin of the second nucleon will oscillate as well).

These forces will derive from an interaction energy:

$$V(r) = -\frac{e^2}{4\pi \mu^2} \left( \mathbf{\sigma}_1 \cdot \mathbf{\sigma}_2 \right) \frac{e^{-kr}}{r}$$  \hspace{1cm} (127)

which serves as a potential energy of this force.

(this is obtained inserting (126) in (118)

Putting now $f = \frac{e^{-kr}}{r}$ one can write (127) in the form:

$$V(r) = -\frac{e^2}{4\pi \mu^2} \left( \frac{\mathbf{\sigma}_1 \cdot \mathbf{\sigma}_2}{r^2} - \frac{\mathbf{\sigma}_1 \cdot \mathbf{\sigma}_2}{3} \right) e^{-kr} \left( f'' - \frac{1}{r} f' \right)$$  \hspace{1cm} (128)

where the terms in the first row represent obviously a tensor force and those in the second can be written by using

$$\nabla^2 f = K^2 f - 4\pi \delta(x),$$

We then get:

$$V(r) = -\frac{e^2}{4\pi \mu^2} \left( \frac{\mathbf{\sigma}_1 \cdot \mathbf{\sigma}_2}{r^2} - \frac{\mathbf{\sigma}_1 \cdot \mathbf{\sigma}_2}{3} \right) e^{-kr} \left( f'' - \frac{1}{r} f' \right)$$  \hspace{1cm} (129)

$$+ \frac{e^2}{\mu^2} \delta(r) \left( \mathbf{\sigma}_1 \cdot \mathbf{\sigma}_2 \right)$$

It is interesting to see what happens if we put $K = 0$, but keep $\mu$ finite (after all, $\frac{e}{\mu}$ is just the coupling constant).

We then get $f'' - \frac{1}{r} f' = \frac{3}{r^2}$. The first and the third term of $V(r)$ gives just the interaction energy of two magnetic dipoles with the strength $\frac{e}{\mu} \mathbf{\sigma}$. The second term vanishes. The third term is a local interaction, repulsive for parallel spin, attractive for anti-parallel one.
This is to be expected since our interaction is just modelled after the interaction of magnetic dipoles, whose fields obviously would have $\kappa = 0$. The second term in (129) is characteristic for a field with a finite mass.

This force is still not satisfactory for the description of nuclear forces since it would vanish in the average over parallel and antiparallel spins. We know, however, that the nuclear forces are attractive in the average. It will be shown that only charged fields give satisfactory nuclear forces.
Fifth lecture

(Notes collected by A. Zichichi)

The Physical significance of strong and weak "couplings"

We would like to measure in some way the strength of interaction of the field with the source. It is related with the strength of the static field, which this source produces.

This strength can be measured as follows: let us consider a source with a static field. Let us imagine the source suddenly being removed. Then the previously static field will spread in all directions since, without source, the field obeys a wave equation. We ask ourselves: how many field-quanta does this spreading field contain after the source was instantaneously suppressed. If this number is less than unity, the coupling is weak; if it is higher than the unity, the coupling is strong. The number of quanta in the static field is given by the following formula:

\[ V(1) \quad N = \sum_{\mathbf{k}} \int \left( F_\mathbf{k} \right)^2 \omega_\mathbf{k} \, d\mathbf{r} \]

where \( F_\mathbf{k} \) is the Fourier component of the wave number \( \mathbf{k} \) of the field and where the integration is carried out over all space. Formula \( V(1) \) can be understood in the following way: we use formula (101) which gives the Fourier analysis of the field in terms of the creation and destruction operators:

\[ V(2) \quad F(\mathbf{x}) = \sum_{\mathbf{k}} F_\mathbf{k} = \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_\mathbf{k}}} \left( C e^{-i\mathbf{k}\cdot\mathbf{r}} + C^* e^{i\mathbf{k}\cdot\mathbf{r}} \right) \]

and put this into \( V(1) \) which gives:

\[ N = \sum_{\mathbf{k}} \left( C C^* + C^* C \right) \frac{\omega_\mathbf{k}}{2\omega_\mathbf{k}} = \sum_{\mathbf{k}} \left( n_{\mathbf{k}} + \frac{1}{2} \right). \]

Hence we got the total number of quanta, if we disregard the contribution \( \sum_{\mathbf{k}} (\xi) \) from the zero point oscillations. Formula
V(1) is evident when we observe that the number of quanta in an oscillator is the square of its amplitude multiplied by the frequency. (Put the mass equal to unity.)

We can also derive formula V(2) in a more familiar way. Let us consider, for instance, the electric field

\[ V(4) \quad E = \sum_k E_k \quad (k \text{ being the index for the different frequencies of the Fourier components of the electric field.}) \]

To see how many quanta there are in the field, we have first to find the total energy belonging to each frequency \( \omega_k \), that is:

\[ V(5) \quad \frac{1}{8\pi} \int \frac{\omega_k^2}{\omega_k} d\tau \]

Then we have to divide it by its proper \( \omega_k \) times \( h \), and to sum over all values of \( k \); the total number of quanta will be:

\[ V(6) \quad N = \sum_k \frac{1}{\omega_k} \int E_k^2 d\tau \quad (h = 1 \text{ in our units}) \]

The difference between V(6) and V(2) is due to the fact that in V(2) we have \( \omega_k E_k^2 \) and in V(6) \( \omega_k E_k^2 \). The origin of this difference lies in the fact that usually one does not use as field quantity the electric field itself, but the potential. For the sake of simplicity we may take a gauge \( E = -A = -i\omega A \) and using \( A \) as the field quantity, V(6) and V(2) become identical.

* * *

**Scalar Field**

Using V(2) we now calculate the number of quanta in the case of scalar coupling. The field is \( F = g \frac{e^{-k \tau}}{r} \). The Fourier components of this field are

\[ F_k = g \frac{1}{k^2 + K^2} \]

and the number of quanta is:
\[ V(7) \quad N = g^2 \sum_k \omega_k \frac{1}{(k^2 + K^2)^2} = g^2 \int \frac{dk^3}{(k^2 + K^2)^2} \omega_k \quad \text{(using } \omega_k = \sqrt{k^2 + K^2} \text{)} \]

\[ = g^2 \int_0^\infty \frac{dk^3}{\omega_k^3} \approx g^2 \log \frac{\overline{K}}{K} \]

(For \( K \ll K \), \( \omega_k \approx K \); so very roughly speaking \( \frac{dk^3}{\omega_k^3} \approx \frac{K^2 dk}{K^3} = \frac{dK}{K} \))

Therefore we can say that the integral is \( \approx \left[ \log \frac{\overline{K}}{K} \right]_K \)

where \( \overline{K} \) is some upper limit of \( K \). Formula \( V(7) \) is of some interest; it shows that \( N \) is proportional to \( g^2 \). It also shows that for \( K \) going to \( \infty \) we have an infinite number of quanta. This is because we have a point-source. Precisely speaking, we always have a strong coupling because the field becomes infinitely strong at the centre. But if we are able to justify the introduction of a finite dimension for the source, for instance if we suppose that the trouble at the centre has something to do with the self-mass and we introduce some kind of finite radius (and therefore a maximum momentum) then the number of quanta is essentially dependent on \( g^2 \) (in fact the ratio enters in the formula through the logarithm). Concluding we have shown that the fact that we have a large or small number of quanta (strong or weak coupling) depends on whether \( g^2 \) is large or small. We now know that if we want to assign \( g \) for the nuclear forces, \( g^2 \) comes out to be \( \frac{1}{10} \). Therefore this would give us the result that the scalar coupling is a relatively weak coupling.

\* \* \*

The Coulomb Field

The same conclusions can be obtained when considering the Coulomb field. In this case:

\[ V(8) \quad N = \frac{g^2}{4\pi} \log \frac{\overline{K}}{K} \]

but now \( K = 0 \) and we have a divergency. This is the so-called "infrared catastrophe". This divergence derives from the fact that the Coulomb field has an infinite range and therefore contains many long wavelength components; clearly when the wavelength becomes very long, \( \pi \omega \) becomes
very small. Therefore, every Coulomb field has a very large number of quanta of low frequency. But this is not a very serious trouble. In fact it can be avoided by saying, for instance, that one never has a Coulomb field reaching to \( \infty \), because electrons are always around to shield it. The shielding of the Coulomb field has the same effect of making \( \kappa \) finite. Also if we introduce a reasonable cut-off at the centre, for example, the electron-radius, the log \( \frac{R}{\kappa} \) becomes some reasonable number, which is not large. The number of quanta is then essentially \( \frac{\alpha^2}{\hbar c} = \frac{1}{137} \), which is a small number. With all these precautions (which are due to the difficulties of Field Theory), we can say that the Coulomb field and the Scalar field are essentially weak coupling fields because it is the coupling constant itself that determines the number of quanta. In other words if you take an electron and spirit it away and look at the field that remains, you have only \( \frac{1}{137} \) as value for the probability of finding one quantum. This means, from the field point of view, that the electro-static field of an electron is essentially a small perturbation. This means that it is a quite good approximation to speak of a "naked-electron", i.e. an electron without Coulomb field. But we must be careful when speaking of an electron "without Coulomb field", because of the important restriction that we have to ignore what happens at the center. Actually the electron has its mass and this mass is probably electromagnetic and comes from its innermost region, where there are strong couplings and one cannot speak of a "naked-electron" any more. We are also allowed to speak of a bare proton in the case of scalar coupling, but we shall see in the next section that this is not true for the ps-coupling.

* * *

The Pseudoscalar Field

As we have seen in the fourth lecture, the ps-field due to a point source is:

\[
\mathbf{V}(9) \quad \mathbf{F} = \frac{\mathbf{\kappa}}{\mu} \left( \nabla \times \mathbf{\varphi} \right) \frac{\mathbf{\varphi}}{r} - \frac{\mathbf{K} \cdot \mathbf{r}}{r^3}
\]

The Fourier components of this field are:
\[ V(10) \quad F_K = \frac{1}{K^2 + K^2} \kappa^2 \]

The corresponding number of quanta is:

\[ V(11) \quad N = \frac{\kappa^2}{\mu^2} \int \frac{dK^3}{\omega_K^3} (\mathbf{\sigma} \cdot \mathbf{\nu}_K)^2 \]

The term \((\mathbf{\sigma} \cdot \mathbf{\nu}_K)^2\) has the magnitude \(\frac{1}{3} K^2\), then we have

\[ N = \frac{\kappa^2}{\mu^2} \cdot \frac{1}{3} \int_K \frac{K^2 K^2 dK}{\omega_K^3} \omega_K \sim K \]

then we obtain:

\[ V(11b) \quad N \approx \frac{\kappa^2}{\mu^2} \frac{R^2 - K^2}{6} \approx \frac{g^2 (\frac{M}{\mu})^2 \cdot \frac{1}{6} \sim 1} \]

The upper limit of \(K\) is \(K = M\) \((M = \text{nuclen mass})\) because we exclude the core of the nucleon, i.e., a region of the order of \(1/M\). The lower limit of \(K\) is not important for \(V(11b)\), in fact it does not appear in the denominator, as in \(V(7)\) but it has to be subtracted from \(K\) and can then be forgotten. It is easy now to see that we cannot any longer say that the bare nucleon is a good approximation, because there is at least one quantum around the proton. The field around the proton is an essential part, not only at the centre, but even in the region outside the "internal" one. In the same region where for the electron the Coulomb field was only a small perturbation there is no chance for the bare-nucleon to be a good approximation.

* * *

Compton-scattering cross-section

In classical physics it is very easy to calculate the scattering of a low-energy light-quantum with a free electron. Let \(E\) be the electric field of the incoming plane waves. We can suppose that this light wave is so weak that any movement of the electron will be so small that we can forget magnetic effects. Under the influence of the electric wave, which has frequency, let us say, \(\omega\), the electron will move up and down. It will perform a harmonic motion that we can calcu-
In general if we have an equation of this kind

\[ (\nabla^2 + k^2) F_\omega = I_\omega \quad (F \text{ and } I \text{ having the same frequency } \omega) \]

the asymptotic solution for the scattered wave (in the direction \( \vec{k}_f \)) coming from this inhomogeneity \( I_\omega \) is:

\[ V(16) \quad \frac{F(\vec{k}_f)}{scattered} = \frac{e^{ik \cdot r}}{r} \int e^{-i\vec{k}_f \cdot \vec{r}'} I(\vec{r}') \, d\vec{r}' \]

In our particular case with the use of \( V(13) \):

\[ V(17) \quad F(\vec{k}_f)_{scat.} = \frac{e^{ik \cdot r}}{r} \frac{e^{i\vec{k}_f \cdot \vec{r}'} 2}{m \omega^2} \int \frac{e^{-i\vec{k}_f \cdot \vec{r}'} \delta(\vec{r}') \cdot \nabla F_0}{m \omega^2} \, d\vec{r}' = \]

\[ = \frac{e^{ik \cdot r}}{r} \frac{1}{m \omega^2} (\vec{k}_f \cdot \vec{k}_o) F_0 \nabla F_0 = i \vec{k}_o F_0 \]

In fact what remains from the integral in \( V(17) \) after partial integration is made, is:

\[ (\vec{k}_f \cdot \vec{k}_o) \delta(\vec{r}') e^{-i\vec{k}_f \cdot \vec{r}'} F_0 \quad \text{and} \quad \delta(\vec{r}') e^{-i\vec{k}_f \cdot \vec{r}'} = 1 \]

The cross-section is the square of this amplitude divided by the incoming field-amplitude squared:

\[ V(18) \quad \sigma = \left( \frac{k^2}{m} \right)^2 \left( \frac{k_f \cdot k_o}{\omega^2} \right)^2 \left[ \frac{k_f \cdot k_o}{\omega^2} \right]^2 = \frac{k^4}{\omega^4} (\eta_o \cdot \vec{k}_f) \Xi \]

Formula \( V(18) \) also shows the angular distribution of the scattered radiation: the strongest intensity will be forward and backward. In other words, such a motion radiates mostly forwards and backwards. A trivial remark: if \( m \to \infty \), i.e. if we fix the particle, the scattering cross-section is zero (see also fourth lecture). The scalar coupling does not give any scattering from an infinitely heavy particle. In fact if the particle is infinitely heavy the incoming field cannot do anything to it. If the particle has, for instance, a spin, then the incoming field can do something to its spin even if the mass is very large and the particle remains at rest.

* * *
where:

\[ X' = X_o + X_p \]

- Position of the particle at rest.
- Displacement of the particle induced by the incident field.
- Position where we measure the field.

Knowing now the induced motion of the source, we have to solve the equation of motion of the field

\[ \nabla^2 F - \ddot{F} - k^2 F = g \delta (x' - x) \]

We have already solved this equation for the static approximation (fourth lecture). Now we no longer consider the particle at rest and we have to see what happens to the right-hand side of \( V(14) \), if the particle moves under the action of an incident plane wave. The particle does not move up and down, like the electron previously, but backwards and forwards because the gradient of \( F \) acts in the longitudinal direction. Now if the particle moves, this motion, on the basis of this inhomogeneity, will produce waves; and we have to calculate these waves which are scattered waves. Now let us forget that \( \delta \) is a strange function and let us suppose that it is some normal function, something like a bell-function. We assume that the field is very weak, then the motion is extremely small and we can write the RHS of \( V(14) \) in the following way:

\[ g \delta (x' - x) = g \delta (x_o) + g \delta' (x') x_p \]

this part produces the static field.

In equation \( V(14) \) we are only interested in terms periodic in \( \omega \). Because \( k^2 = \omega^2 - \kappa^2 \) equation \( V(14) \) becomes:

\[ \nabla^2 F + \kappa^2 F = g \delta' (x') x_p \]
In general if we have an equation of this kind

$$(\nabla^2 + \kappa^2) F_\omega = I_\omega \quad (F \text{ and } I \text{ having the same frequency } \omega)$$

the asymptotic solution for the scattered wave (in the direction $\vec{k}_f$) coming from this inhomogeneity $I_\omega$ is:

$$V(16) \quad F_{\text{scattered}}^{(k_f)} = \frac{e^{ik \cdot r}}{r^3} \int e^{-i \vec{k}_f \cdot \vec{r}'} I(r') \, dr'^3$$

In our particular case with the use of $V(13)$:

$$V(17) \quad F_{\text{scat.}}^{(k_f)} = \frac{e^{ik \cdot r}}{r^3} \int \frac{e^{-i \vec{k}_f \cdot \vec{r}'} \delta'(r')}{\omega^2} \nabla F_0 \, dr'^3 =$$

$$= \frac{e^{ik \cdot r}}{r^3} \frac{e^{2 \omega}}{m \omega^2} (\vec{k}_f \cdot \vec{k}_0) F_0 \nabla F_0 = i \vec{k}_f F_0$$

In fact what remains from the integral in $V(17)$ after partial integration is made, is:

$$(\vec{k}_f \cdot \vec{k}_0) \delta(r') e^{-i \vec{k}_f \cdot \vec{r}} F_0 \quad \text{and} \quad \delta(r') e^{-i \vec{k}_f \cdot \vec{r}} = 1$$

The cross-section is the square of this amplitude divided by the incoming field-amplitude squared:

$$V(18) \quad \sigma = \left( \frac{e^2}{m} \right)^2 (\frac{\vec{k}_f \cdot \vec{k}_0}{\omega^2})^2 \left[ \frac{(\vec{k}_f \cdot \vec{k}_0)^2}{\omega^2} = \frac{k_f^4}{\omega^4} (\vec{\eta}_0 \cdot \vec{\eta}_f) \right]$$

Formula $V(18)$ also shows the angular distribution of the scattered radiation: the strongest intensity will be forward and backward. In other words, such a motion radiates mostly forwards and backwards. A trivial remark: if $m \to \infty$, i.e. if we fix the particle, the scattering cross-section is zero (see also fourth lecture). The scalar coupling does not give any scattering from an infinitely heavy particle. In fact if the particle is infinitely heavy the incoming field cannot do anything to it. If the particle has, for instance, a spin, then the incoming field can do something to its spin even if the mass is very large and the particle remains at rest.
Scattering of pseudoscalar pions

As we have seen in the fourth lecture, the spin equation of motion due to a pseudoscalar field is:

$$V(19) \quad \sigma_1 = \frac{E}{\mu} (\mathbf{\sigma}_o \times \mathbf{\nu})$$

It gives the change in time of \( \sigma \) when there is a field \( \mathbf{\nu} \).

To compare the situation with the scalar case we remember that for the scalar coupling the effect of the field is given by \( \mathbf{m} \mathbf{x} = - e \mathbf{\nabla} \mathbf{F} \). But now we make the assumption that the mass of the heavy particle tends to infinity, because we are only interested in the spin effects and we forget the motion of the particle. We assume that \( \mathbf{\nu} \) has a time dependence \( e^{i \omega t} \), then \( \mathbf{\nu} = i \omega \mathbf{\nu} \), and from \( V(19) \) we get:

$$V(20) \quad \sigma_1 = \frac{E}{\omega \mu} (\mathbf{\sigma}_o \times \mathbf{k}_o) \mathbf{\nu}_o \quad \nabla \mathbf{\nu} = - i \mathbf{k}_o \mathbf{\nu}$$

\( \mathbf{k}_o \) is the incoming wave vector.

Formula \( V(20) \) is very interesting. It shows that \( \sigma_1 \) is perpendicular to \( \sigma_o \); \( \sigma_o \) being the spin of the heavy particle at rest. Then the actual \( \sigma \) will be:

$$\sigma = \sigma_o + \sigma_1$$

\( |\sigma_1| \) is very small, compared with \( |\sigma_o| \), and has time dependence \( e^{i \omega t} \).

This does not have time dependence.

Now we have to see what radiation is produced by such a vibration of the spin around \( \mathbf{\sigma}_o \). Through the influence of these vibrations the equation of motion of the field \( \mathbf{\nu} \) becomes according to \( (124) \):

$$V(21) \quad (\nabla^2 - K^2) \mathbf{\nu} - \mathbf{\nu} = - \frac{E}{\mu} (\sigma_1 \times \mathbf{\nu}) \mathbf{\delta}(x)$$

(we omit the static part which does not lead to radiation). The scattered wave given by this inhomogeneity \( \sigma \) is:

$$V(22) \quad \mathbf{\eta}_{\text{scat.}} = \frac{e^{i k \mathbf{r} \cdot \mathbf{\sigma}_1}}{\mu} \int e^{-i k \mathbf{r} \cdot \mathbf{\nu}} (\sigma_1 \times \mathbf{\nu}) \mathbf{\delta}(x - x') \, dr'^3$$
By partial integration of \( V(22) \) and replacing \( \sigma_1 \) by \( V(20) \) we have:

\[
\mathcal{E}_{\text{scatt}} = \frac{4\pi r^2}{\mu^2 \omega} \left( \mathbf{\hat{r}}_0 \times \mathbf{\hat{r}}_f \right) \cdot \mathbf{\hat{r}}_f \mathbf{\hat{r}}_0 \left( \mathbf{\hat{r}}_0 \times \mathbf{\hat{r}}_f \right) \cdot \mathbf{\hat{r}}_0
\]

the cross-section will be:

\[
\sigma_{\text{scatt.}} = \left[ \frac{2}{\mu^2 \omega} \left( \mathbf{\hat{r}}_0 \times \mathbf{\hat{r}}_f \right) \cdot \mathbf{\hat{r}}_0 \right]^2
\]

which is like the cross-section for magnetic dipole radiation.

\[
* \quad * \quad *
\]

**Final remarks**

We know that it is not correct to say that \( \mathbf{\sigma}_o \) is constant in time. \( \mathbf{\sigma}_o \) has its own frequency (only the z-component is constant and the x-y component oscillate). Then \( \mathbf{\sigma}_1 \) does not only have the frequency \( \omega \) of the incoming field, as we have assumed. The fact that \( \mathbf{\sigma}_o \) has its own frequency has the consequence that the inhomogeneity in \( V(21) \) has also an additional term and more properly it would be

\[
V(23) = \frac{\mathcal{E}}{\mu} (\mathbf{\sigma}_1 \times \mathbf{\nabla}) + \frac{\mathcal{E}}{\mu} (\mathbf{\sigma}_o \times \mathbf{\nabla})
\]

We have omitted the second term because we have considered \( \mathbf{\sigma}_o \) static. But actually it is not static; therefore, the inhomogeneity \( V(23) \) has also a time-dependent term with its own frequency \( \omega_n \).

Now the problem arise: if \( \mathbf{\sigma}_o \) has its internal vibration, why does it not radiate? Classically it would, of course. But quantum mechanically not. The situation is similar to that of the H-atom which has internal frequencies on its ground state, but does not radiate. The internal frequency \( \omega_n \) only adds terms to the scattering cross-section which were neglected in our calculation.
Charged Fields

Until now we have considered only uncharged meson fields. As long as the interactions with nucleons are neglected the classical free Hamiltonian has the form

\[ H = \frac{1}{2} \int \left( -\mathbf{p}^2 + (\nabla F)^2 + \kappa F^2 \right) \, d\tau \]

where \( F \) is real.

If we want that in the expression for the total energy (VI-1) of our field \( F \), a contribution should arise from the electromagnetic field, we have to introduce it in the \( H \) (VI-1). If we express the e.m. field as the 4-vector \((V, \mathbf{A})\) of the e.m. potential, then to introduce the electromagnetic energy in the expression for the total energy (VI-1) we have to make the following well known substitutions:

\[ \frac{\partial}{\partial x} \rightarrow \frac{\partial}{\partial x} - ieA \]

VI (2)

\[ \frac{\partial}{\partial t} \rightarrow \frac{\partial}{\partial t} + ieV \]

With these changes (VI-1) becomes the Hamiltonian of the meson-field when the contribution of the e.m. field is also taken into account. An immediate consequence of this change (VI-2) is that
F can no longer be a real quantity. F must become a complex field. A complex field effectively means that every field has two components

\[ F = F_1 + i F_2 \]  

where \( F_1 \) and \( F_2 \) are real fields. The fact that \( F \) is no longer a "real" quantity introduces a change in the Hamiltonian (VI-1) which is intuitive. In fact we now have to replace \( F^2 \) by the square of the modulus of \( F \) and so on; (VI-1) then becomes

\[ H = \frac{1}{2} \int \left( \left| \hat{\mathbf{F}} \right|^2 + \left| \nabla F \right|^2 + \kappa^2 \left| F \right|^2 \right) \, d \mathcal{V} \]  

It is important to appreciate the differences between (VI-1) and (VI-4).

If \( F \) in (VI-1) is replaced by any complex field, \( F_1 + i F_2 \), where \( F_1 \) and \( F_2 \) are independent and not correlated by the relations (VI-2), then (VI-4) represents the sum of the two separate Hamiltonians of \( F_1 \) and \( F_2 \).

But, the introduction of the electromagnetic field into the Hamiltonian in our physical problem leads to a correlation between the two quantities \( F_1 \) and \( F_2 \). This correlation arises from the substitutions (VI-?), that is from gauge-invariance. Under these essential conditions, (VI-4) is no longer a simple sum since terms involving \( F_1 \) and \( F_2 \) occur in the real and imaginary parts and cannot be separated.
We now calculate the current $\mathcal{J}$ and the charge density $\mathcal{P}$ of the field.

The current $\mathcal{J}$ and the charge density $\mathcal{P}$ are directly given once we know the Hamiltonian; in fact

\[ \mathcal{J} = \frac{\partial H}{\partial A(x)} \]

\[ \mathcal{P} = -\frac{\partial H}{\partial V(x)} \]

which are explicitly expressed by

\[ \mathcal{J} = \frac{ie}{2} \left[ F^*(\nabla - isA)F - F(\nabla + isA)F^* \right] \]

\[ \mathcal{P} = \frac{ie}{2} \left[ F^*(\frac{\partial}{\partial t} + isV)F - F(\frac{\partial}{\partial t} - isV)F^* \right] \]

It is most interesting to look at the particles far from the region of interaction (when $A = V = 0$), because in this case the expressions become

\[ \mathcal{J} = \frac{ie}{2} \left( F^* \nabla F - F \nabla F^* \right) \]

[where $\pi = \frac{\partial L}{\partial F}$]

\[ \mathcal{P} = \frac{ie}{2} \left( F^* \pi - \pi F \right) \]

[then $\frac{\partial F}{\partial t} = \pi$]

In terms of $F_1$ and $F_2$ we can express $\mathcal{J}$ and $\mathcal{P}$ as follows:
\[ J = \frac{e}{2} (F_1 \nabla F_2 - F_2 \nabla F_1) \]

VI (9)
\[ \mathcal{J} = \frac{e}{2} (\pi_1 \pi_2 - \pi_2 \pi_1) \]

We can now easily see that \( F \) must be complex in order to have a non-zero charge, because if either \( F_1 \) or \( F_2 \) is zero, both \( J \) and \( \mathcal{J} \) vanish. If \( F \) is purely real or purely imaginary, we have no charge. One can also easily see that the charge will depend on the sign of the frequency \( \omega \). In other words, if we have a plane wave

VI (10) \[ F = Ae^{i(Kx - \omega t)} \]

the charge will be defined by the sign of \( \omega \). In fact, the charge density of this wave will be

VI (11) \[ \mathcal{J} = \pm |A|^2 e^{i \omega} \]

and the sign \( \pm \) is determined by the time dependent term in (VI-10) as can easily be understood by looking at the expression for \( \mathcal{J} \) (VI-7).

The total charge will be

\[ \mathcal{J}_{\text{tot}} = \int \mathcal{J} d\mathbf{r}^3 = \pm |A|^2 e^{i \omega} \]

Again if the incoming wave is represented by a real quantity

\[ F = \cos (Kx - \omega t) \]

\( \mathcal{J} \) will obviously be zero.
We can then see how the use of a complex field has physical significance and is important.

In order to find the significance of the number \( e \) we have to know what is \( 'A' \); as we know, the amplitude of the wave which has one quantum is \( \frac{1}{\sqrt{\omega}} \). This comes from the expression of the field operator in terms of the destruction and construction operators. Then if we substitute this value of \( A \) in (VI-11) we find for the total charge

\[
\mathcal{Q}_{\text{tot}} = \mp e
\]

If we calculate the current for the same plane wave (VI-10) we have

\[
\mathcal{J} = \pm |A|^2 e \hat{F}
\]

and putting in it the amplitude for one quantum we have

\[
\mathcal{J} = \frac{\hbar}{\omega} e = \mathcal{J} e
\]

(in fact: \( \frac{\hbar}{\omega} = \frac{\text{momentum}}{\text{energy}} = \text{velocity} \). These are the results that we had to expect once we quantized our field. In fact, the total charge of a plane wave containing only one quantum cannot be different from \( \pm \) the elementary charge \( e \). And its current is just given by \( \mathcal{J} e \).

* 

*  

*
Final Remarks Concerning :

1) The Pseudoscalar Interaction Between Two Nucleons

On the basis of such an interaction we have calculated (see 4th Lecture, page 30-129) the interaction energy between two nucleons. We have also said that the force we get is almost analogous to that between two dipoles. Now we want to emphasize the difference contained in this quasi-analogy. As everybody knows, the dipole forces average out to zero, and the nuclear forces do not average out to zero. We want to point out that if in (129) we put $K = 0$ the second term vanishes and the analogy with dipole forces is complete. But if we keep $K$ finite, as it must be for a field with finite mass, then the second term remains and the interaction energy does not average out to zero any more.

2) The Scattering of Pseudoscalar-Pions

In the Final Remarks of the last lecture (V) on treating the problem of scattering of pseudoscalar-π, we said that we made one approximation; namely to consider $\phi_0$ constant in time. We now want to add that apart from this approximation made within the framework of quantum theory, we have also made another approxima-
tion, this time in the framework of classical theory. Namely, it is not true to say that only the incoming field acts upon the spin. This approximation had as a consequence that $\phi_0$ was used in (V-20), where $\phi_0$ is the incoming field. In fact, during the radiation there is the radiated field which also reacts on the nucleon and, as a result, we can no longer consider $\phi_0$ as the only field acting on the particle.
3) The Introduction of Complex "Field" or "Wave-Function" in Quantum Theory

This is something different from the introduction of complex wave-function in, for example, the theory of alternating currents. In the theory of alternating currents only the real part has any physical significance and the rest is just a trick to make calculations more simple. But, if you like, at the expense of complicated equations, you could write everything with only the real parts. In quantum theory, by contrast, both the real and imaginary parts of the wave-function have physical significance; there are many expressions that need both. And in a way you can say that the Schroedinger equation is a two-component wave-function which connects real and imaginary parts.

In the theory of charged fields the fact that both the two fields are physically significant will become more pronounced.
Quantization of charged fields (Pauli and Weisskopf, H.P.A., 709 (1934))

In the previous lecture it has been shown that the description of charged quanta requires a complex field $F(X)$. The complex nature of $F(X)$ was seen to result from the substitution $\frac{\partial}{\partial x^\mu} - \frac{\partial}{\partial x^\mu}$, by which the Lagrangian describing motion under the influence of electromagnetic fields is obtained from the Lagrangian of the free motion. From the definition of $\rho$ and $\mathcal{J}$ (equ. (1) and (2)) it is seen that the signs of the charge density and the current density depend on the time dependance of $F(X)$. For example for a plane wave in the positive $k$-direction:

$$F = e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$  corresponds to positive charge.

$$F = e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$  "  negative charge.

$$F = \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)$$  "  no charge.

In order to quantize a charged field, we split the classical field $F(X)$ into the real and the imaginary parts.

VII(1)  \[ F(X) = F_1(X) + iF_2(X) \]

($F_1$ and $F_2$ are real functions satisfying the classical equation of motion (46)). In analogy with (58) we expand $F_1$ and $F_2$ in terms of normal vibrations $f_k(X)$:

VII(2)  \[ F_1(X) = \sum_k q_1^{(k)} f_k(X) \quad (i = 1, 2) \]

and obtain:

VII(3)  \[ F(X) = \sum_k q^{(k)} f_k(X) \]

4614
where:

\[ q(k) = q_1(k) + iq_2(k) \]

As pointed out in the second and the third lecture, the \( q_i(k) \) are real quantities, i.e. they become hermitean operators if we pass from classical to quantum theory. It follows then from VII (4) that the \( q(k) \) are non-hermitean operators, since they contain hermitean and anti-hermitean parts. Since we are not so much interested in the real and the imaginary parts of \( F(x) \), but rather in the parts corresponding to annihilation and creation of quanta, we write, in analogy with (93):

\[ q_1(k) = \frac{1}{\sqrt{2\omega_k}} \left( c_1(k) + c_1(k)^* \right) \]

creation operator

anihilator operator

For reasons of simplicity we drop the index in the following formulas, keeping in mind that we have to sum over \( k \) in the final expressions for \( F, \Pi, \rho, \) etc. Furthermore, we define a new set of operators:

\[ a = \frac{1}{\sqrt{\omega}} \left( c_1 + ic_2 \right) \]
\[ a^* = \frac{1}{\sqrt{\omega}} \left( c_1^* - ic_2^* \right) \]
\[ b = \frac{1}{\sqrt{\omega}} \left( c_1 - ic_2 \right) \]
\[ b^* = \frac{1}{\sqrt{\omega}} \left( c_1^* + ic_2^* \right) \]

(Note that \( a \) and \( b \) are not complex conjugates of each other because \( c_1 \) and \( c_2 \) are complex!)

With the new quantities we get instead of VII (5):

\[ q = \frac{1}{\sqrt{\omega}} \left( a + b^* \right) \]
\[ q^* = \frac{1}{\sqrt{\omega}} \left( a^* + b \right) \]

In order to get expressions for \( \rho, F, H \), etc. in terms of \( a, a^*, b \) and \( b^* \), we must transform our commutation rules (39) for the \( c \)'s.

From (39) and VII (6), one finds:
VII (8) \[ [a, a^*] = [b, b^*] = 1 \]
\[ [a, a] - [b, b] = 0 \]

Thus we see that the transformation VII (6) from the \( C_1 \), \( C_2 \) to the 
\( a, b \) is a "canonical" one: the \( a, b \) have the same commutation rules
as the \( C_1, C_2 \). The significance of the operators VII (6) can be illustrated as follows: we know that \( C_1 \) and \( C_2 \) are destruction operators,
\( C_1^* \) and \( C_2^* \) are creation operators. More specific:

Say, \( C_1^* \) and \( C_2^* \) create their real fields, each having a time dependence \( \cos(\omega t + \delta) \). Then the expression VII (6) for \( a^* \) shows that \( a^* \) creates both a quantum of the field \( F \), and a quantum of the field \( F^*_2 \)
with a phase difference of \( \frac{-\pi}{2} \), since \(-i = e^{\frac{-i\pi}{2}}\). Hence the created
field \( F = F_1 + iF_2 \) has a time dependence \( \cos(\omega t + \delta) - i \sin(\omega t + \delta) = \)
\( = e^{-i\omega t} \cdot e^{i\delta} \) which corresponds to positive charge. In the same way
we would have made plausible that \( b^* \) creates a field of negative charge; from this would follow, of course, that \( a \) destroys a positive charge and \( b \) destroys a negative one.

We can prove that these assignments are correct by finding the operator of total charge in terms of these quantities. For the fields \( F \) and \( F^* \) and their momentum conjugates \( \Pi = \hat{F}^* \) and \( \Pi^* = \hat{F} \) we obtain:

VII (9) \[ F = \sum_k \frac{1}{\sqrt{\omega_k}} (a_k + b_k^*) f_k \]
\[ F^* = \sum_k \frac{1}{\sqrt{\omega_k}} (a_k^* + b_k) \bar{f_k} \]

VII (10) \[ \tau = i \sum_k \sqrt{\omega_k} (c_k^* - c_k) f_k \]
\[ \bar{\tau} = i \sum_k \sqrt{\omega_k} (-c_k - c_k^*) \bar{f_k} \]

Inserting VII (9) and VII (10) into the expression ( ) for \( \rho \), and using the commutation rules VII (6) we find:

VII (11) \[ Q = \int \rho \, d^3x = e \sum_k (a_k^* a_k - b_k^* b_k) = \]
\[ = e \left\{ \left( \text{number of pos. quanta} \right) - \left( \text{number of neg. quanta} \right) \right\} \]
as one expects. From this we can easily compute the commutators of
Q with any of the operators VII (6). For example:

VII (12) \[ [Q, a_k] = - e a_k \]

which means indeed, that the operator \(a_k\) acting on a state \(\Psi\) with
charge \(q'\) produces a state with charge \(q' - e\), as one can see in the
following way: the fact that \(q'\) is the charge of the state \(\Psi\) is
expressed by: \(Q \Psi = q' \Psi\). Operate on this equation from the left
with \(a_k\): \((a_k Q) \Psi = q'a_k \Psi\) \((q'\) is a number and commutes with \(a_k\)). From
VII (12) follows that \((a_k Q) \Psi = (q'a_k) \Psi + e a_k \Psi\) and hence
\(Q(a_k \Psi) = (q' - e)(a_k \Psi)\) which means that the charge of \(a_k \Psi\) is \((q' - e)\).

Since \(a_k\) is a destruction operator, it must correspond
to the destruction of a positive quantum, as we have conjectured above.
In a similar way the character of \(a^*\), b and \(b^*\) could be demonstratable.

It is important to observe from VII (9) that the field
operator \(F\) always creates a positive charge or created a negative
one, whereas the field operator \(F^*\) always creates a positive charge and de-
stroues a negative one. Hence \(\Psi\) decreases the total charge and \(F^*\) in-
creases it by one unit.

Next we express the hamiltonian \(H\) in terms of
VII (6). In the same way as in the case of the charge operator, we
find:

VII (13) \[ H = \sum_k \omega_k \left( a_k^* a_k + b_k^* b_k \right) \]

i.e. the total energy is equal to the sum of the energies of the posi-
tive and the negative quanta plus the corresponding zero point ener-
gies. The absence of a zero point term in the expression for the total
charge VII (11) results from the cancelation of the charge fluctuation
+ \(\frac{e}{2}\) and \(- \frac{e}{2}\) from the positive and the negative quanta of each energy \(\omega_k\).
On the other hand, the zero point energy of a state is positive, no
matter whether the quanta have positive or negative charge. Hence we
get a contribution of one quantum \(\omega_k\) to each term in the sum of VII(13).
It must be reminded that the above derivations are based on the assumption, that the quanta, inspite of being charged, do not interact appreciably with one another. As soon as there are interactions between the bosons the hamiltonian ( ) can no longer be expanded in terms of free harmonic oscillators, because in addition to the term $|\nabla F|^2$, there will be coupling terms due to $\varphi$ and $\bar{\varphi}$. We will assume in the following, that such terms are small.

*  
*  
*  

Interaction of a charged field with nuclecus

a) Scalar interaction.

There is a characteristic difference between the interaction of charged and uncharged fields. For an uncharged field there was no restriction on the number of quanta (mesons) which may be emitted or absorbed by the source (nucleon). In the case of a charged field the requirement that the charge must be conserved limits the number of emitted or absorbed quanta. Therefore the interaction term

\begin{equation}
\text{VII (14)} \quad J = \int g F(X) \delta (X-q) \, d^3X
\end{equation}

used in the scalar theory of neutral mesons may not be applied to the theory of charged mesons. We have seen above that $F(X)$ decreases the positive charge of the meson field. In order to conserve charge we must introduce an additional operator which increases the charge of the nucleon by the same amount, such that the overall charge is conserved. Since the only way to increase the charge of a nucleon is the conversion of a neutron into a proton, we need an operator with the following properties: it yield a proton wave function when applied to a neutron wave function, and it gives zero when applied to a proton wave function. From nuclear physics we know that the isotopic spin operator $\mathbf{T}_- = \frac{i}{2} (\mathbf{T}_1 - i\mathbf{T}_2)$ has just this property. Therefore, from the point of view of charge conservation the replacement $F(X) \rightarrow F(X) \mathbf{T}_-$ would yield a suitable coupling term.
However, this term violates time reversal invariance; it only allows for decrease of the meson charge and increase of the nucleon charge and not the opposite transition! The formalist would say that, because of the complex nature of $F(X)$ such a coupling term is not hermitean and a non hermitean hamiltonian is a violation of time reversal. We therefore must add another term of equal strength for the opposite process:

\[ J = g \int (F_\tau^* + F^* \tau) \delta (x - q) \, d^3 x \]

which now satisfies charge conservation and time reversal. The latter fact is exhibited formally by the reality of $J$. (Note that $\tau^* = \tau$.)

However, this interaction does not contain the possibility of emission and absorption of a neutral meson. In order to take this into account we must add a new real field component $F^\tau_3$ to our field, which now consists of a complex field $F = F^\tau_1 + iF^\tau_2$ and an additional real field $F^\tau_3$. One also can speak of 3 real field components: $F^\tau_1, F^\tau_2, F^\tau_3$. This real field is coupled to the proton and to the neutron by a coupling of the type VII (14) as it is natural for real (and therefore uncharged) fields. In order to get the utmost symmetry between the three fields one chooses the coupling constant $g_3$ for the third field thus by: $g_3 = g$ for protons and $g_3 = -g$ for neutrons. One then can write the interaction as follows:

\[ J = g \int \left\{ F_\tau^* + F^* \tau^+_\tau + F^\tau_3 \tau_3 \right\} \delta (x - q) \, d^3 x \]

where $\tau_3$ is the Pauli-matrix \[ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \] which is unity for a proton wave function and minus one for a neutron. This form is of particular symmetry in a vector-terminology in the so called isotopic-spin space: consider $F^\tau_1, F^\tau_2, F^\tau_3$ as the three components of a vector in the space, and also the three Pauli-operators $\tau_1, \tau_2, \tau_3$. Then:

\[ F_\tau^* + F^* \tau^+_\tau + F^\tau_3 \tau_3 = F^\tau_1 \tau_1 + F^\tau_2 \tau_2 + F^\tau_3 \tau_3 = (\mathbf{F} \cdot \mathbf{\tau}) \]

where the vector symbol below indicates a vector in the iso-space. Hence we get:

\[ J = g \int (\mathbf{F} \cdot \mathbf{\tau}) \delta (x - q) \, d^3 x \]
as the most symmetric scalar interaction of three types of meson
(+, −, uncharged) with a nucleon. It is so called charge-independent
form in which the field $F_1$, $F_2$, $F_3$ represents a vector in the isoto-
pic space, which corresponds to an iso-spin unity for the meson field,
just as a vector field in the ordinary space corresponds to an ordi-
nary spin unity (see light quanta).

b) Spinor interaction.

The interaction VII (19) is not yet the one which is
realized in nature because of the absence of the spin of the nucleon.
Also the field $F$ would be a scalar in the ordinary space whereas we
know it to be a pseudoscalar. As before in the case of a neutral field
we therefore introduce a $(\sigma \cdot \nabla)$ - coupling. Before we had the expres-
sion :

$$J = \frac{f}{\mu} \int (\sigma \cdot \nabla) \tilde{\epsilon} \delta (x - q) \, d^3x$$

calling the field $\tilde{\epsilon}$ now, instead of $F$. This can easily be generalized
to a 3-isotopic component field $\tilde{\epsilon}_1 \tilde{\epsilon}_2 \tilde{\epsilon}_3$ by writing :

$$VII \quad J = \frac{f}{\mu} \int (\sigma \cdot \nabla) (\tilde{\epsilon}_1 \tau^+ + \tilde{\epsilon}_2 \tau^+ + \tilde{\epsilon}_3 \tau^+) \delta (x - q) \, d^3x$$

$$\quad \quad = \frac{f}{\mu} \int (\sigma \cdot \nabla) (\tilde{\epsilon}_\lambda \tau_\lambda) \delta (x - q) \, d^3x$$

an expression which contains vector products in the real space and
in the isospin space.
Static Nuclear Force in the Charge Independent Spinor Theory

The occurrence of $\mathcal{Z}$ in the interaction term means that in addition to the space coordinates $q$ and the spin $\sigma$, the nucleon has an additional coordinate. This corresponds to the well-known fact that neutrons and protons can convert themselves into each other. We know that formally $\mathcal{Z}$ behaves very much like a mechanical angular momentum operator, i.e., it obeys the same commutation rules as $\sigma$. This fact will be made use of later.

With the introduction of an interaction we change two things. We change the motion of the nucleon and we change the motion of the field. Let us start with the field. As we have seen in the fourth lecture, the equation of motion (46) of the free field goes over into an equation with the left side unchanged, but with the zero on the right side replaced by a source function. The source function is obtained by observing that $J$ enters with minus sign into the Lagrangian, and by taking its derivative with respect to the components of $\mathcal{Z}$. One then finds:

$$VIII \ (1) \quad -\nabla^2 \phi + m^2 \phi + \mathcal{Z} \phi = \frac{\kappa}{\mu} (\vec{\sigma} \cdot \vec{\nabla}) \phi (x-q) \mathcal{Z},$$

where we have written the three equations for $\phi_1$, $\phi_2$ and $\phi_3$ as a vector equation in isospacce. This equation gives us the field around the source. If we now ask for a static solution of VIII (1) we again encounter the same difficulty as in the case of the uncharged field, namely that only the $z$-component of $\vec{\sigma}$ is fixed, whereas the other two will oscillate and will only be zero in the time average. The same is true for the isotopic spin. So all we can do is to calculate the time average of the field by putting
\( \phi = 0 \), and assuming \( M \to \infty \), so the nucleon is fixed in space, say at \( q = 0 \). In the same way as before we find:

\[
\text{VIII (2)} \quad \phi_{\text{static}} = \frac{-e}{4\pi \mu} (\vec{\sigma} \cdot \vec{\nabla}) \frac{e^{-Kr}}{r} \quad \text{(time average)}
\]

So we have a result quite similar to (126), the only difference being that here \( \phi \) is a vector in isospace, whereas before it was a scalar in isospace. We can use VIII (2) to compute the potential energy between two nucleons which are a certain distance \( r \) apart from each other, by inserting \( \phi_{\text{static}}(r) \) of VIII (2), i.e. the field produced by nucleon 1 at the position \( r \) of nucleon 2, into the interaction term of nucleon 2. We must be aware that the result obtained will be valid only for large distances, because in that case the interference terms due to the simultaneous fluctuations of \( \vec{\zeta}_1 \) and \( \vec{\zeta}_2 \), and \( \vec{\sigma}_1 \) and \( \vec{\sigma}_2 \) will be incoherent (independent) and vanish in the average, i.e. the spins and the isospins of both nucleons may be considered fixed.* The result is:

\[
\text{VIII (3)} \quad V_{12}(r) = -\frac{e^2}{4\pi \mu^2} (\vec{\sigma}_1 \cdot \vec{\nabla}_1) (\vec{\sigma}_2 \cdot \vec{\nabla}_2) e^{-Kr} \frac{1}{r} (\vec{\zeta}_1 \cdot \vec{\zeta}_2)
\]

which differs from (128) of lecture IV only in the occurrence of the scalar product of the two isovectors \( \vec{\zeta}_1 \) and \( \vec{\zeta}_2 \). As before, we can split the potential into three parts

\[
\text{VIII (4)} \quad V_{12}(r) = \frac{e^2}{4\pi \mu^2} \left[ 2 \left( \left( \frac{1}{2} \sigma \cdot r^2 \right) \vec{\zeta}_1 \cdot \vec{\zeta}_2 \right) + \frac{1}{3} (\vec{\sigma}_1 \cdot \vec{\sigma}_2) (\vec{\zeta}_1 \cdot \vec{\zeta}_2) K^2 f \right] \left[ \sigma \cdot \vec{r}_1 \sigma \cdot \vec{r}_2 \right] \cdot (\vec{\zeta}_1 \cdot \vec{\zeta}_2)
\]

where

\[
S = \frac{(\vec{\sigma}_1 \cdot \vec{r})(\vec{\sigma}_2 \cdot \vec{r})}{r^2} - \frac{1}{3} (\vec{\sigma}_1 \cdot \vec{\sigma}_2)
\]

is the tensor operator

\[
\sigma = \frac{e^{-Kr}}{r}
\]

* To give some meaning to this procedure, we may think of an experiment where two well collimated beams of nucleons are passing each other at a distance of the order of several Compton wavelengths. These beams will be deflected by their mutual attraction. It is the potential giving rise to this deflection that we compute.
The first term is the tensor force which contributes very little to nuclear binding. The second term is a Yukawa potential multiplied by the scalar products of the spin operators and the isospin operators of the two nucleons. It is easy to show that this term gives attraction for nucleons very close together, i.e. in S-states. We recall that

\[
(\vec{\sigma}_1 \cdot \vec{\sigma}_2) = \begin{cases} 
-3 & \text{for spin } S = 0 \text{ (singlet)} \\
+1 & \text{" } S = 1 \text{ (triplet)}
\end{cases}
\]

and similarly

\[
(\vec{\tau}_1 \cdot \vec{\tau}_2) = \begin{cases} 
-3 & \text{for isospin } T = 0 \text{ (isosinglet)} \\
+1 & \text{“ } T = 1 \text{ (isotriplet)}
\end{cases}
\]

For two nucleons in an S-state the Pauli principle requires:

\[
T = 1 \text{ for } S = 0 \\
T = 0 \text{ for } S = 1
\]

so the Yukawa function in the second term of VIII (4) is always multiplied by a minus sign, i.e. this state is always attractive in S-states. It gives the main contribution to nuclear binding.

In the same way we see that the third term has a plus sign for S-states, i.e. it is repulsive for nucleons very close together. This term corresponds to the repulsive core. (As we have seen in Fubini's lecture, the quantum mechanical treatment of VII (20) requires the \(\delta\)-function in the integrant to be replaced by a source function of finite extension. Such a replacement obviously also results in a finite range of the third term of eq. VIII (4).)

It must be remembered, however, that VIII (4) was derived from the assumption of a rather large distance, so in any case we cannot trust our results at small distances. This difficulty of computing the nuclear force at small distances is also encountered in the quantum mechanical treatment of this problem.
There is one point in which the quantum mechanical treatment is superior to the classical one used here. In the above formulas we have always made the silent assumption that we are dealing with bare particles, i.e. we have neglected the effects of the meson cloud. However, the operators $\sigma$ and $\zeta$ used in our formulas refer to the real nucleons. For example, the eigenvalue of $\zeta_3$ is +1 for a real proton and -1 for a real neutron. But part of the time a proton consists of a neutron plus a $\pi^+$-meson, so the bare particle of a proton is not an eigenstate of $\zeta_3$. Similar considerations apply to the spin $\sigma$: because the virtual mesons are emitted in p-states, the spin of the bare particle can flip, while $\sigma_3$ will still remain sharp for the physical nucleon. As has been shown in the Fubini lecture, the taking into account of these virtual effects results in a reduction of the coupling constant. So our formulas are correct if we replace $f$ by $f(\text{renormalized})$. In the following we will simply write $f$ instead of $f(\text{renormalized})$.

* *

* *

**Scattering of $\pi$-mesons on nucleons**

As previously done in the neutral case (lecture VI), the scattering amplitude will be computed on a completely classical basis. An incoming plane wave representing the meson beam is scattered by a nucleon, which is fixed in space. The only nucleon coordinates that may change are $\vec{\sigma}$ and $\zeta$. So the only effect of the incoming wave on the nucleon will be to induce flips of spin and of isospin. We will assume that the interaction is such that these flips very seldom occur. In the classical picture this corresponds to saying that changes of the angular momentum vectors $\vec{\sigma}$ and $\zeta$, induced by the incoming wave, are very small. So the actual values of $\vec{\sigma}$ and $\zeta$ will be given by
\[ \sigma^2 = \sigma^2_o + \sigma^2_1(\omega) \]
\[ \zeta = \zeta^2_o + \zeta^2_1(\omega) \]

where \( \sigma^2_o \) and \( \zeta^2_o \) are the average values of \( \sigma^2 \) and \( \zeta \), and \( \sigma^2_1 \) and \( \zeta^2_1 \) are small changes induced by the incoming wave, the phase factor of which we assume to be

\[ \psi_o \sim e^{i(k_o \cdot r - \omega t)} \]

As in lecture V, our over-all \( \psi \) is composed of the incoming part plus the scattered part

\[ \psi = \psi_o + \psi_{\text{scattered}} \]

and \( \psi_{\text{scattered}} \) is given by

\[ \psi_{\text{scattered}} = \frac{e^{iKr}}{4\pi r} \int e^{-iK_1 \cdot r} S(q, \sigma, \zeta) d^3r \]

where \( S \) is the source function, i.e. the inhomogeneity term of eq. VIII (1). In our classical approach to this problem the scattering is due to the fluctuating part of \( S \), so we have to insert

\[ \psi(\zeta, \sigma) - \psi_o(\zeta_o, \sigma_o) = \int \left\{ (\sigma_1 \cdot \nabla) \zeta_2 + (\sigma_2 \cdot \nabla) \zeta_1 + (\sigma_1 \cdot \nabla) \zeta_1 \right\} \delta(x-q) d^3r \]

rather than \( S(\zeta, \sigma) \) into VIII (8) to get \( \psi_{\text{scattered}} \).

Since the last term may be neglected compared to the others, we see that \( \psi_{\text{scattered}} \) will be composed of a part corresponding to the
fluctuations of ordinary spin plus a part corresponding to the fluctuations of isospin. In lectures IV and V we have seen that the time derivative of $\sigma$ may be obtained by considering the analogy between our interaction VII (20) and the interaction term of a magnetic dipole moving in a magnetic field (eqs. 119 - 122). \( \omega_d \), (122) may, of course, also be obtained from the quantum mechanical equation.

\[
\mathbf{\dot{\sigma}} = \frac{1}{i} [\mathbf{\sigma}, \mathbf{J}]
\]

and an analogous equation holds for $\mathbf{\dot{\zeta}}$. Using VII (5) and VII (7), and keeping only the largest terms, one finds:

\[
\mathbf{\dot{\zeta}}_1 = \frac{2e}{\mu} (\mathbf{\sigma}_0 \times \mathbf{V}) (\mathbf{\rho}_0 \cdot \mathbf{\zeta}_0)
\]

\[
\mathbf{\dot{\zeta}} = -\frac{2e}{\mu} (\mathbf{\sigma}_0 \cdot \mathbf{V}) (\mathbf{\rho}_0 \times \mathbf{\zeta}_0)
\]

(where the minus sign in the last equation results from the interchange of the order of $\mathbf{\rho}_0$ and $\mathbf{\zeta}_0$ in the cross product as compared to $\mathbf{\sigma}_0$ and $\mathbf{\rho}_0$, in the first equation). Because of VII (6) the differential operations on both sides are readily carried out, yielding:

\[
\mathbf{\dot{\sigma}}_1 = \frac{2e}{\mu} (\mathbf{\sigma}_0 \times \mathbf{\zeta}_0) (\mathbf{\rho}_0 \cdot \mathbf{\zeta}_0)
\]

\[
\mathbf{\dot{\zeta}}_1 = \frac{2e}{\mu} (\mathbf{\sigma}_0 \cdot \mathbf{\zeta}_0) (\mathbf{\rho}_0 \times \mathbf{\zeta}_0)
\]

By use of VII (9) and (12) we obtain for the scattered wave VII (8)
VIII (13) \[ \psi_{\text{scatt}} = \frac{2i\hbar^2}{\mu^2} \left\{ (\vec{\sigma}_0 \times \vec{K}_0) \vec{K}_f (\vec{\omega}_0 \times \vec{\omega}_0) \psi_0 + (\vec{\sigma}_0 \cdot \vec{K}_0) (\vec{\sigma}_0 \cdot \vec{K}_f) (\vec{\omega}_0 \times \vec{\omega}_0) \right\} \frac{e^{iK'r}}{4\pi r} \]

This expression may be rewritten by the use of the following formulas:

VIII (14) \[ (\vec{\omega}_0 \times \vec{\omega}_0) \psi_0 = \psi_0 + i(\vec{\omega}_0 \times \vec{\omega}_0) \]

\[ (\vec{\sigma}_0 \cdot \vec{K}_f)(\vec{\sigma}_0 \cdot \vec{K}_0) = (\vec{K}_f \cdot \vec{K}_0) + i \vec{\sigma}(\vec{K}_0 \times \vec{K}_f) \]

One finds:

VIII (15) \[ \psi_{\text{scatt}} = \frac{-2i\hbar^2}{\mu^2 \omega} \left\{ (\vec{K}_0 \times \vec{K}_f) \vec{\sigma}_0 \psi_0 + (\vec{K}_0 \cdot \vec{K}_f) (\vec{\omega}_0 \times \vec{\omega}_0) \right\} \frac{e^{iK'r}}{4\pi r} \]

In the next lecture we will show that this result is identical with that of the quantum mechanical treatment in the Born approximation.
Ninth lecture

(Notes collected by G. Weber)

Some Supplementary Remarks on Charge Independence

In lecture VII we have introduced the so-called charge independent form of the spinor interaction, in which the meson field consists of real fields $F_1$, $F_2$, and $F_3$, which represent a vector in charge space and therefore correspond to isotopic spin unity for the mesons. The charge independence is considered a consequence of the fact that in VII(20) the field is combined with the isospin operator $\tau$ to a scalar product in isospace, i.e., an invariant in isospase. Therefore, once we define charge independence in this way, there is no reason, why we should not add to VII(20) a term containing another real field $g_0 F_0$. Since $F_0$ is a scalar in isospace, it would correspond to neutral pions with isotopic spin zero. The existence of such a neutral pion with $T = 0$ has recently been postulated in theoretical papers by Yamaguchi and by Baldin. The interaction in this case is (we use the abbreviation $g = f/\mu$ in this section):

$$J = \int \left\{ g \left( \vec{\sigma} \cdot \vec{\tau} \right) (\vec{u} \cdot \vec{v}) + g_0 \left( \vec{\sigma} \times \vec{v} \right) \delta (x - q) \right\} d^3 x$$

The static nuclear force derived from this interaction is given by the sum of the expressions (129) and VII(4), i.e., it contains terms essentially proportional to $(\vec{\sigma}_1 \cdot \vec{\sigma}_2)$ $(\vec{T}_1 \cdot \vec{T}_2)$ and terms proportional to $(\vec{\sigma}_1 \times \vec{\sigma}_2)$. It is not clear, whether the introduction of these latter terms is consistent with nuclear data. Checking the above expression against nuclear physical data is very difficult, because at small distances, where the nuclear forces are strong, the expressions (129) and VII(4) are no longer applicable.

From what we have said so far, it seems rather arbitrary that we introduced in our interaction just 4 real fields $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$ and $\vec{u}_0$ and not any larger number of fields. There is a good reason for
using 4 fields only as we can see as follows: we usually describe the nucleon as a 2-component isospinor with the neutron and proton as two orthogonal eigenstates:

\[
\begin{align*}
\psi_p &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \psi(x_1 \sigma) = \begin{pmatrix} p \\ 0 \end{pmatrix} \\
\psi_n &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \psi(x_1 \sigma) = \begin{pmatrix} 0 \\ n \end{pmatrix}
\end{align*}
\]

and as a consequence of this, the Hamiltonian is a $2 \times 2$ matrix acting on the isospinor part of the wave function. We recall that any $2 \times 2$ matrix may be represented as a linear combination of the isotopic spin matrices $\tau_1, \tau_2, \tau_3$ and the unit matrix $\tau_0$. (Please note that IX(1) is such a linear combination, because the last term acts on neutrons the same way as on protons, i.e. it may be thought of as multiplied by $\tau_0$, and the first term is a linear combination of $\tau_1, \tau_2, \tau_3$). The most general form of interaction may obviously be written in the form (we omit here all factors and operators which are irrelevant for the present consideration):

\[
J \sim \sigma_1 F_1 \tau_1 + \sigma_2 F_2 \tau_2 + \sigma_3 F_3 \tau_3 + \sigma_0 F_0 \tau_0
\]

with arbitrary $\sigma_i$'s. If we require the Hamiltonian to obey certain physical conservation laws, we get restrictions for the values of $\sigma_i$. For example, we have seen in lecture VII that charge conservation and time reversal invariance require $\tau_1$ and $\tau_2$ to occur in the form:

\[
g(F \tau_+ + F^* \tau_-) = g(F_1 \tau_1 + F_2 \tau_2)\] which gives $\sigma_1 = \sigma_2$.

We have also seen that charge independence requires $\sigma_1 = \sigma_2 = \sigma_3$. Let us derive this relation. Charge independence implies that we can choose, instead of IX(2) any two orthonormal combinations of $\psi_p$ and $\psi_n$ without changing the Hamiltonian. For example, if we choose instead of

\[
\begin{pmatrix} p \\ n \end{pmatrix} \text{ a new pair of states } \begin{pmatrix} p' \\ n' \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} p + n \\ p - n \end{pmatrix}
\]

these states $p'$ and $n'$ cannot be realized in nature, because they represent particles which are neither protons nor neutrons. Nevertheless, if charge independence holds, the laws of physics expressed in terms
of these states must be the same as before. To prove this, we observe that:

\[ \left( \begin{array}{c} p' \\ n' \end{array} \right) = \frac{\tau_1 + \tau_3}{\sqrt{2}} \left( \begin{array}{c} p \\ n \end{array} \right) \]

i.e. we have just made a unitary transformation

\[ \psi' = U \psi; \quad \psi'^* = \psi^* U^* \]

and therefore, we must transform the operators accordingly:

\[ \tau'_1 = U \tau_1 U^*, \quad \tau'_2 = -U \tau_2 U^* \]

Carrying out IX(6) gives for the new operators:

\[ \tau'_1 = \tau_3, \quad \tau'_2 = -\tau_2, \quad \tau'_3 = \tau_1. \]

This obviously is a rotation in charge space, i.e. the choice of a new coordinate system. In order not to change physics, we must transform the field in the same way:

\[ F'_1 = F_3, \quad F'_2 = -F_2, \quad F'_3 = F_1. \]

Instead of IX(3), we get:

\[ J \sim \sigma_1 F_3 \tau_3 + \sigma_2 F_2 \tau_2 + \sigma_3 F_1 \tau_1 + \sigma_0 F_0 \tau_0 \]

which is identical with IX(3) for \( \sigma_1 = \sigma_3 \). By similar transformations we could derive \( \sigma_1 = \sigma_2 = \sigma_3 \neq \sigma \). The term with \( \sigma_0 \) is not affected by any of these transformations and therefore \( \sigma_0 \) may be different from \( \sigma \). In lectures VII and VIII we have assumed \( \sigma_0 = 0 \), as is usually done in current theories.

* * *

**Meson-Nucleon Scattering in Born Approximation**

We want to prove that the expression VII(15) for the scattering, which we derived in a classical way, is identical with
that of the Born approximation calculation. To prove this, we have to rewrite \( \text{VIII}(15) \) in terms of a scattering amplitude. We ask for the amplitude of the scattered wave in the direction \( k_f \). If we denote by \( \mathbf{g}_o \) the amplitude of a wave describing one quantum moving in direction \( k_f \), then the amplitude of the scattered wave corresponding to one quantum in the incoming beam is given by:

\[
\text{IX}(9) \quad (\mathbf{g}_o, \mathbf{g}_{\text{scatt}}) = \frac{-2e^2 \text{i}}{\mu^2 \omega} \left[ (\mathbf{P}_o \times \mathbf{P}_f) \mathbf{a}_o (\mathbf{g}_o \cdot \mathbf{g}_f) + \right. \\
+ (\mathbf{P}_o \cdot \mathbf{P}_f) \mathbf{a}_f (\mathbf{g}_o \times \mathbf{g}_f) \right]
\]

In quantum mechanics we proceed as follows: we have at \( t = -\infty \) a state \( i \) containing \( \{1 \text{ meson } \mathbf{P}_o, \mathbf{g}_o, \text{1 nucleon} \} \) and a similar state \( f \) at \( t = +\infty \) containing \( \{1 \text{ meson } \mathbf{P}_f, \mathbf{g}_f, \text{1 nucleon} \} \) and we ask for the transition probability between states \( i \) and \( f \). According to the "Golden Rule", this probability is given by:

\[
\text{IX}(10) \quad T = \text{const} \times |M|^2
\]

and our problem is essentially that of computing \( M \). In first order perturbation theory, we would just have to take the matrix element of the interaction term \( \text{VII}(20) \) between the two states:

\[
\text{IX}(11) \quad M = \langle i | J | f \rangle \quad (1\text{st order})
\]

However, the interaction \( \text{VII}(20) \) allows for the emission or the absorption of one meson at a time only, and such an emission or absorption act certainly is not a scattering event. This matrix element therefore does not occur in our problem. So we go to the second order approximation:

\[
\text{IX}(12) \quad M = \sum_n \frac{\langle i | J | n \rangle \langle n | J | f \rangle}{E_i - E_n} \quad (2\text{nd order})
\]

The intermediate states \( n \) are well defined. There are 2 types of such states as can be seen by illustrating the process by Feynman diagrams:
In fig. 1, the intermediate state has 0 mesons, 1 nucleon.
In fig. 2, " " " " 2 " 1 "

Our sum IX(12) consists out of two terms, each resulting from one type of diagrams. If we fix the energy scale such that the free nucleon has zero energy, than the energy denominator belonging to fig. 1 is $\omega$, the one belonging to fig. 2 is $-\omega$. We therefore obtain:

$$
IX(13) \quad M = \frac{f^2}{\mu^2 \omega} \left\{ \sum_n <n|(\vec{\sigma} \cdot \vec{K}_o)(\vec{q}_o \cdot \vec{\tau})|n>|n|(\vec{\sigma} \cdot \vec{K}_f)(\vec{q}_f \cdot \vec{\tau})|f> - \sum_n <n|(\vec{\sigma} \cdot \vec{K}_f)(\vec{q}_f \cdot \vec{\tau})|n>|n|(\vec{\sigma} \cdot \vec{K}_o)(\vec{q}_o \cdot \vec{\tau})|f> \right\}
$$

For each value of $\omega$, there are four states $n$ (corresponding to 2 possible orientations of spin and of isospin of the nucleon). Therefore we can replace $\sum_n |n><n|$ by unity:

$$
IX(14) \quad M = \frac{2f^2}{\omega \mu^2} \left\{ |n|(\vec{\sigma} \cdot \vec{K}_o)(\vec{q}_o \cdot \vec{\tau})|n> - (\vec{\sigma} \cdot \vec{K}_f)(\vec{q}_f \cdot \vec{\tau})|n> \right\}
$$

This may be rewritten by means of VIII(14):

$$
IX(15) \quad M = \frac{2f^2}{\mu^2 \omega} \left\{ |n|(\vec{K}_o \cdot \vec{K}_f)(\vec{q}_o \cdot \vec{q}_f) \tau + \vec{\sigma} \cdot (\vec{K}_o \times \vec{K}_f)(\vec{q}_o \cdot \vec{q}_f) |f> \right\}
$$

This result is essentially the same as IX(8), which was obtained on a completely classical basis. (The difference in sign results from the imaginary character of $M$ and the fact that IX(8) represents $M^*$ rather than $M$.)
The Quantization of Fermi Field

We are going to look at the Fermi particles. The Fermi particles are ordinary particles like electrons, protons and neutrons, whereas the Bose particles are the quanta which we obtain as the result of the quantization of classical fields: photons for the electromagnetic field and mesons for the nuclear field.

In the classical limit, the Fermi particles are "real" particles: the equation of motion of these particles is described by the classical Hamiltonian. The quantization of this Hamiltonian brings about some wave properties. On the other hand, in the Bose case, we start with classical fields which acquire particle properties through the quantization. In other words, after the quantization, the classical particles and fields are much more similar to each other. Hence the quantized equations exhibit symmetrical properties between the Fermi particles and Bose fields and lead to a lot of similarities, some of which are almost identical, except for the commutation relations in which the sign is positive for the Fermi particles instead of negative for the Bose particles.

Now, in order to quantize the Fermi particles, we have to start with the quantum mechanics of one particle. Later, we shall go on to the quantum mechanics of many particles; and at the very end we will have to deal with an infinite number of fermions, because according to the Dirac positron theory, the vacuum is essentially an infinite number of particles in negative
energy states. However, at present we will concentrate our attention on cases of finite number of particles. The characteristic feature of the fields of fermions is the conservation of the number of particles. We know that the equations are made in such a way that the number of fermions is constant.

Let us look at the Hamiltonian of one single particle:

\[ X(1) \quad H = H(p,q) \]

and the operators

\[ X(2) \quad p \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial q}, \quad q \rightarrow q, \quad H \rightarrow -\frac{\hbar}{i} \frac{\partial}{\partial t}, \]

then we get the wave equation:

\[ X(3) \quad H(p,q) \psi(q,t) = -\frac{\hbar}{i} \frac{\partial}{\partial t} \psi(q,t), \]

and the energy eigenvalue equation is:

\[ X(4) \quad H(p,q) \psi_a(q) = E_a \psi_a(q), \]

where \( p \) and \( q \) are just operators and we determine the Schrödinger wave function as a function of \( q \). \( E_a \) is the eigenvalue corresponding to the stationary state with certain index \( a \).

We are now working in the configuration space of coordinate \( q \), then the wave function \( \psi \) is described as a function of \( q \). However, we can work in the polar coordinate system or we can also work in the momentum space in which we have different assignments for the operators, namely:

\[ X(5) \quad p \rightarrow p, \quad q \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial p}. \]
Then we get the differential equation in the coordinate $p$, and it will automatically give us the wave function which depends on $p$ instead of $q$. In general, the operator $p$ and $q$ must fulfil the relation

$$X(6) \quad \left[ p, q \right] = i\hbar$$

and if two operators fulfil this relation, we can choose these as operators for $p$ and $q$.

The transformation theory allows us to calculate the wave function in one space if the wave function is known in another space. As a matter of fact, the $p$-wave function is just the Fourier transform of the $q$-wave function. In this theory, the Hamiltonian need not be the free particle Hamiltonian, but may be any external field acting on the particle. However, we will not go any further into the transformation theory.

One important fact is that to every operator ($p$, $q$, energy, or angular momentum) there exists a set of wave functions in which the magnitude is sharp. In the configuration space, the wave functions in which the momentum is sharp will be given as the plane waves:

$$X(7) \quad e^{ip_a q} \rightarrow p = p_a$$

and the eigenfunction of $q$ in the configuration space will be:

$$X(8) \quad \delta(q - q_a) \rightarrow q = q_a .$$

From the eigenvalue equation:

$$X(9) \quad q \delta(q - q_a) = q_a \delta(q - q_a)$$

we see that, for the function $X(8)$, the eigenvalue of $q$ is $q_a$. 

4950/E
In the momentum space, the situation is completely the opposite. In the momentum space, the wave functions in which \( q \) is sharp are the plane waves:

\[
X(10) \quad e^{i q_a p} \rightarrow q = q_a
\]

and the wave functions in which \( p \) is sharp must be the Fourier transforms of \( X(7) \) and are:

\[
X(11) \quad \delta(p - p_a) \rightarrow p = p_a.
\]

We see that in the momentum space the \( p \)-function is the \( \delta \) function. Obviously, in the \( y \)-space, the function in which \( y \) is sharp must be the \( \delta \) function in \( y \).

**Many Identical Particles Without Interaction**

We shall go back to the discussion of the situation of many particles without interaction among themselves. This might be a many-electron problem: for example, the system of electrons in the uranium atom without interaction between them. However, we are keeping the number of particles finite.

In order to construct the Schrödinger equation for many fermions, we begin with the Hamiltonian of a many-particle system without interaction:

\[
X(12) \quad H_o = \sum_{\alpha=1}^{N} h_\alpha (p_\alpha, q_\alpha),
\]

where \( h_\alpha(p,q) \) is the one-particle Hamiltonian and \( \alpha \) is the particle number. There is no interaction term in the Hamiltonian. Our wave function of the many-particle system \( \psi_A \) must obey the Schrödinger equation:

\[
X(13) \quad H_o \psi_A = E_A \psi_A
\]
where $\psi_A$ is the eigenfunction for the total system and $E_a$ is the eigenvalue. We can then write down:

$$X(14) \quad \psi_A(q_1, q_2, \ldots, q_N) = \varphi_a(q_1) \varphi_b(q_2) \ldots \varphi_z(q_N),$$

where $q_1, q_2, \ldots, q_N$ are the coordinates for each particle and $\varphi'$s are the solution of:

$$X(15) \quad h_a(p, q) \varphi_a = \epsilon_a \varphi_a(q), \ldots$$

with $\epsilon_a, \ldots, \epsilon_z$ being corresponding eigenvalues for each state.

If the particles are independent (i.e. no interaction), the quantum mechanical solution must be a product such as $X(14)$. The eigenvalue of $\psi_A$ is:

$$X(16) \quad E_A = \sum_a \epsilon_a$$

Here we see that we can have $N!$ solutions for $\psi_A$ having the same eigenvalue $E_a$, since we can permute the particles among the states which they occupy. Here enters the Pauli principle. We can use any linear combination of the eigenfunctions for the solution of our problem. However, if these particles are identical and indistinguishable, it then becomes necessary to make a decision between many linear combinations.

Now, Pauli saw from nature that the wave function must have one special symmetry, namely antisymmetry: in other words, although we can make any linear combinations for our problem, nature selects the antisymmetric ones.

According to Pauli, we take the following wave function:

$$X(17) \quad \psi_A = \frac{1}{\sqrt{N!}} \left[ \varphi_a^{(1)}, \varphi_b^{(2)}, \ldots, \varphi_z^{(N)} \right]$$

$$+ \sum \pi \varphi_a^{(1)}, \varphi_b^{(2)}, \ldots, \varphi_z^{(N)} \right]$$
where \( \sqrt{\frac{1}{N!}} \) is the normalization factor, each of \( a, b, \ldots, z \), assigns the states of one of the particles, \( \mathcal{P} \) denotes the permutation between the particles in different states, and \( \Sigma \) represents the summation of all the possible permutations. Then we can split \( \Psi \) into the two groups, namely even permutation and odd permutation, together with the sign + and -:

\[
X(18) \quad \psi_A = \sqrt{\frac{1}{N!}} \left[ \sum \pi^e \varphi_a(1), \ldots \varphi_z(N) - \sum \pi^o \varphi_a(1), \ldots \varphi_z(N) \right].
\]

The first term of \( X(17) \) is included in the even term of \( X(18) \). Now the antisymmetry has the following effect:

**None of the states are equal:**

\[
X(19) \quad a \neq b \neq c \neq, \ldots, \neq z
\]

because if one or more states were equal, the interchange of the two states must alter the sign without changing the wave function. Hence \( \psi_A = 0 \). We can also see that

**No two particles have the same position:**

\[
X(20) \quad q_1 \neq q_2 \neq q_3 \neq, \ldots, \neq q_N
\]

because we can choose the states in such a way that the eigenvalue for each \( \varphi \) is the position. Hence, no two equal particles can be at the same position. This fact is closely connected with the impenetrability in the classical sense that no two particles can occupy the same place at the same time. Actually, in the quantum mechanics, the minimum distance between each particle depends very much upon the relative momentum between these particles. In order to clarify this situation, let us take a system of two free particles. The wave function of the system is, according to \( X(18) \):
\( \psi(1,2) = \frac{1}{\sqrt{2}} (e^{i\vec{p}_a \vec{x}_1} e^{i\vec{p}_b \vec{x}_2} - e^{-i\vec{p}_b \vec{x}_1} e^{i\vec{p}_a \vec{x}_2}) \)

where each particle is described by a plane wave and \( \sqrt{\frac{1}{M}} \) is \( \sqrt{\frac{1}{2}} \).

Their momenta and coordinates are \( \vec{p}_a \), \( \vec{p}_b \) and \( \vec{x}_1 \), \( \vec{x}_2 \). Hence we can rewrite the wave function:

\[
(1,2) = \frac{1}{\sqrt{2}} \left[ \exp \left( i\frac{(\vec{p}_a - \vec{p}_b) \cdot (\vec{x}_1 - \vec{x}_2)}{2} \right) + \exp \left( -i\frac{(\vec{p}_a - \vec{p}_b) \cdot (\vec{x}_1 - \vec{x}_2)}{2} \right) \right]
\]

Disregarding the factor two, we see that \( \vec{p}_a - \vec{p}_b \), \( \vec{x}_1 - \vec{x}_2 \) are the relative momentum and the relative coordinate, whereas \( \vec{p}_a + \vec{p}_b \), \( \vec{x}_1 + \vec{x}_2 \) are the total momentum and the total coordinate. Taking the centre of mass system and putting \( \vec{p}_a + \vec{p}_b = \vec{p}, \vec{x}_1 + \vec{x}_2 = \vec{x} \), \( \vec{p}_a - \vec{p}_b = \vec{p}, \frac{\vec{x}_1 - \vec{x}_2}{2} = \vec{x} \), it becomes:

\[
\psi(1,2) = \frac{i\vec{p} \vec{x}}{2} \sin \left( \frac{\vec{p} \vec{x}}{2} \right)
\]

This function is antisymmetric as was expected, since changing the sign of \( x \), \( \sin (-x) \) becomes \( -\sin (x) \). We know that here \( |\psi|^2 \) must be the probability for finding the two particles at the relative distance \( x \), so that \( \psi \) is a function of \( x \):

\[
|\psi|^2 = \sin^2 \left( \frac{x}{\lambda} \right)
\]

with:

\[
\lambda = \frac{h}{p}
\]

It builds up a reasonable probability at the distance from the point at which the two particles coincide, as shown in
Fig. 1. Here the momentum p is exactly well-defined since no force acts on them and p must be constant over time.

Actually, the wave function is a superposition of various momenta which is somewhere around p, so that what we actually get is the superposition of different momenta, as shown in Fig. 2.

We now consider the solution of a general problem of many-particles with interactions. The Hamiltonian can be written in the form:

\[ H = H_0 + H' \]

where \( H' \) contains the interactions and \( H_0 \), being the part without interactions, has the form \( X(12) \). The solution could be very complicated and might contain the effect of the Coulomb field or other fields. However, we can always construct the wave function from a set of orthonormal functions which are antisymmetric. (Obviously, the linear combination of antisymmetric function is also antisymmetric.) It is easiest to use the solution of the Hamiltonian \( H_0 \) as our orthonormal set. We then get:

\[ \psi = \sum_A b_A \psi_A \]

where \( \psi_A \) is the solution of \( X(13) \) and \( b_A \)'s are some coefficients.
The index $A$ indicates a given solution of $X(13)$, that is a set of occupied states:

$$X(27) \quad A = A(a,b,c,\ldots z), \quad B = B(a;b;c;\ldots z') \ldots$$

The functions $\Psi_A, \Psi_B \ldots$ are orthonormal to each other so that we can build any other antisymmetric wave function of $n$-particles from the complete set of the $\Psi_A$ as indicated in $X(26)$.

* * *
Eleventh Lecture

It has been shown in lecture X that the wave function of any system with many identical particles can be expanded in terms of product wave function:

\[ \Psi = \sum A \varphi_A, \]

where \( \varphi_A \) is the antisymmetrized product of orthonormal one-particle wave function \( \varphi_a \) as shown in X(13):

\[ \varphi_A = \left( \varphi_a(1), \varphi_b(2), \ldots \varphi_z(N) \right) \equiv [a, b, c, \ldots, z], \]

where \([a, b, \ldots, z]\) is the short-hand expression of the function \( \varphi_A \). We fix the order of states \( a, b, c, \ldots, z \) in a certain natural order - for example, the order of energy: from the lowest energy to higher ones. Interchanging the states, the sign of XI(2) will be either positive or negative, corresponding to an even or an odd permutation as it is antisymmetric. Using the above short-hand expression, XI(1) can be written:

\[ \Psi = \sum_{a, b, \ldots, z} b_{a, b, \ldots, z} [a, b, \ldots, z] \]

The coefficient \( b_{a, b, \ldots, z} \) can be determined by the orthonormal condition, as follows:

\[ b_{a, b, \ldots, z} = \langle \varphi_A^*, \Psi \rangle = \langle [a, b, \ldots, z]^*, \Psi \rangle. \]
The Creation and Destruction Operators for Fermi Particles

In this section we will see how the creation and destruction operators behave for the Fermi particles. Let us suppose that we have a product wave function $\psi_A$ which has a certain order:

\[ a, b, c, \ldots, z \]

If the order is not natural, we can rearrange it to the natural order by an even or an odd permutation together with positive or negative sign, for example:

\[ a, b, c, \ldots, z \rightarrow [a, b, d, c, \ldots, z] \]

Now we define that the operator $C_k^*$ creates the particle at the state $k$ which is supposed to be the last state, namely:

\[ C_k^* [a, b, \ldots, z] = [a, b, \ldots, z, k] \]

In general, the $k$ is not the last one in the natural order, so we can rearrange it to the natural order by an even or an odd permutation:

\[ [a, b, \ldots, z, k] = \pm [a, b, \ldots, k, \ldots, z] \]

If the $\psi_A$ on which we operate does not contain the $k$, the above statement will hold. However, if it already contains the $k$, it becomes $0$ by the Pauli principle. By the same analogy, the destruction operator $C_k$ destroys the particle in the state $k$ which is supposed to be put at the end of the order:

\[ C_k [a, b, \ldots, z, k] = [a, b, \ldots, z] \]
If the k is not the last one in the original order, we can re-arrange it to the last term by an even or an odd permutation:

\[ [a, b, \ldots, k, \ldots, z] = \pm [a, b, \ldots, z, k]. \]

Then the operator can act according to XI(9). If the original state does not contain the k, it becomes 0.

Since we are dealing with Fermi particles, the number of particles must be constant. In any actual process, there will be an equal number of creation and destruction operators.

**Anticommutation Relations**

By defining the creation and destruction operators for our state, we can derive the anticommutation relations. The first one is:

\[ C^*_k \mathcal{L} = - \mathcal{L} C^*_k \]

To prove this, let us apply the operators on \( \Psi_A \), then we can see:

\[ C^*_k \mathcal{L} [a, b, \ldots, z] = [a, b, \ldots, z, \mathcal{L}, k] \]

\[ \mathcal{L} C^*_k [a, b, \ldots, z] = [a, b, \ldots, z, k, \mathcal{L}] \]

Hence:

\[ C^*_k \mathcal{L} = - \mathcal{L} C^*_k \]

If \( \Psi_A \) contains either k or \( \mathcal{L} \), both sides of the equation will be 0, therefore the relation holds in general.
Secondly, we can see:

\[ X(13) \quad C_k^e C_e = - C_e^e C_k \]

because, keeping in mind XI (8),

\[
\begin{align*}
[a, b, \ldots, k, e, \ldots, z] &= \mp [a, b, \ldots, k, e] \\
[a, b, \ldots, k, e, \ldots, z] &= \pm [a, b, \ldots, e, k]
\end{align*}
\]

and

\[
\begin{align*}
C_k^e C_e [a, b, \ldots, z, k, e] &= [a, b, \ldots, z] \\
C_e^e C_k [a, b, \ldots, z, e, k] &= [a, b, \ldots, z]
\end{align*}
\]

In this case also, if the original state does not contain either \( k \) or \( e \), both sides of the equations will be 0. Then this relation holds in general.

Thirdly, we can prove:

\[ X(14) \quad C_k^e C_e^e = - C_e^e C_k^e \quad (k \neq e) \]

because:

\[
\begin{align*}
C_k^e C_e [a, b, \ldots, z, e] &= [a, b, \ldots, z, k] \\
C_e^e C_k [a, b, \ldots, z, e] &= - [a, b, \ldots, z, k]
\end{align*}
\]

If the original state does not contain \( e \) or if it contains \( k \), both sides of XI(15) will be 0. Hence the relation is valid in general. In addition, we see:

\[ X(16) \quad C_k^e C_k^* + C_k^* C_k = I \quad (I : \text{unit matrix}) \]
since if our state does not contain $k$, the first term makes it 0, whereas the second term does not make any change, and if our state contains $k$ these operators act in the opposite way. Hence we proved the relation.

Combining XI (11), (13), (15) and (17), we can write the anticommutation relation as follows:

$$
\left[ C_k^*, C_k \right]_+ = 0
$$

XI (18)

$$
\left[ C_k C \right]_+ = 0
$$

$$
\left[ C_k^* C \right]_+ = \delta_k \ell
$$

where the bracket $[ \ ]$ is just the ordinary commutator bracket except that here the sign is positive.

So far we have defined what the operators $C_k$, $C_k^*$ do to wave functions of the type $[a, b, \ldots, z]$. Since all wave functions of a system of equivalent particles can be built up as linear combinations of such types (see XI(1)), we have defined the $C$'s completely. Evidently, the operators $C_k^*$, $C_k$ depend upon what choice we have made for the orthonormal one-particle functions $\varphi_k(x)$. If we choose a different orthonormal set $\varphi'_k(x)$, we will get different operators $C_k^*, C_k^*$. $C_k^*$ creates a particle in the state $\varphi_k(x)$ while $C_k$ creates it in the state $\varphi'_k(x)$.

We are now going to introduce new operators which we call:

$$
\psi^*(x) = \sum_k C_k^* \varphi_k(x)
$$

XI (19)

and

$$
\psi(x) = \sum_k C_k \varphi_k(x)
$$

(The symbol $\psi(x)$ is unfortunate, but generally used. To indicate that $\psi$'s are not wave functions but operators, we underline them.)
We will now show that this operator does not depend on the choice of the set \( \varphi_k(x) \). In other words, we will show that:

\[
\psi(x) = \sum_k C_k \varphi_k(x) = \sum_k C^*_k \varphi^*_k(x).
\]

Before we proceed to the proof, let us first explain what this operator is. It is a linear combination of \( C \)'s, each with the coefficient \( \varphi_k(x) \). What does a linear combination of \( C \)'s mean? Let us look at the simplest one: the operator \( f_\alpha^a C^*_a + f_\beta^b C^*_b \). This operator creates a particle in the state \( f_\alpha^a \varphi_a + f_\beta^b \varphi_b \). (Of course \( |f_\alpha|^2 + |f_\beta|^2 = 1 \).) The operator \( f_\alpha^a C^*_a + f_\beta^b C^*_b \) annihilates a particle from that state. Hence \( \psi^*(x_o) \) creates a particle into the state \( \sum_k \varphi^*_k(x_o) \varphi_k(x) \) which is the state given by the delta function \( \delta(x-x_o) \) since the following relation holds for any orthonormal system:

\[
\delta(x-x_o) = \sum_k \varphi^*_k(x_o) \varphi_k(x)
\]

Let us prove the same things in a different way. We consider the transformation between the system of \( \varphi(x) \) and that of the other orthonormal system of \( \varphi'(x) \). This can be done by the unitary transformation:

\[
\varphi'^*_k(x) = \sum \varphi^*_k \varphi'^*_k(x)
\]

\[
\varphi'_k(x) = \sum \varphi^*_k \varphi'_k(x)
\]

with

\[
\sum_k \varphi^*_k \varphi^*_{k,m} = \delta_{m,n},
\]

which is the necessary and sufficient condition for the unitary transformation. Then the creation operator has to fulfil the same transformation relation since one creates a particle in \( \varphi_k \) and the other in \( \varphi'_k \).
XI (24) \[ C_k^* = \sum_{l} \alpha_{lk}^* C_l^* . \]

From this follows that the destruction operator fulfills the relation:

XI (25) \[ C_k^* = \sum_{l} \alpha_{lk}^* C_l^* . \]

Then we can write down XI(20) as follows:

\[ \psi'(x) = \sum_k C_k^* \phi_k^* = \sum_k \sum_{l} \alpha_{lk}^* C_l^* \phi_{nk} \phi_n \]

\[ = \sum_l C_l \phi_l = \psi(x). \]

Here we see that \( \psi(x) \) is invariant whatever orthonormal system of \( \phi(x) \) we choose and actually the choice of orthonormal system is quite arbitrary.

Now we can choose the orthonormal system of \( \phi(x) \) in a rather unusual way, namely, the delta function:

XI (27) \[ \phi_a(x) = \delta(x-a). \]

This is also an orthonormal set of functions and the functions of the position. (We can take the one-dimensional space without losing the generality.) It is obvious from X(9) that:

XI (28) \[ x\delta(x-a) = a\delta(x-a), \]

where \( a \) is the eigenvalue which can have all possible values. Logically, if the particle is in the \( a \)-state, it means that this particle is at the position \( a \).

If we choose this system for our wave function, XI(19) becomes:

XI (29) \[ \psi^*(x) = \sum_a C_a^* \delta(x-a). \]
Similarly:

\[ \psi(x) = \sum a^* c_a \delta(x-a). \]

The significance of these equations is this: the right-hand side of them are 0 for \( x \neq a \). Hence all that is left is the creation or annihilation operator: \( c^*_a \) or \( c_a \), that is the creation operator which adds a particle in the \( a \)-state at the position \( a \) or the annihilation operator which subtracts the particle in the \( a \)-state at the position \( a \). We can also write this as follows:

\[ \psi^*(x) = c^*_x, \quad \psi(x) = c_x. \]

These are the creation and destruction operators at the position \( x \).

Let us point out an important fact. We have chosen our wave function as a function of \( q \). However, we can choose our wave function as a function of \( p \), energy, angular momentum or any other operators depending upon what space we are using. As an example, we can use the wave function \( \varphi(p) \) in the momentum space instead of \( \varphi(q) \) in the coordinate space. Then the new operator will be:

\[ \psi(p) = \sum_k c_k \varphi_k(p), \]

which is different from the operator \( \psi(q) \) because \( \varphi(p) \) is not the same number as \( \varphi(q) \). However, the operator \( c_k \) will be the same, since \( \varphi_k(p) \) is only another representation of the same state as \( \varphi_k(x) \). The question is now, what is the significance of the operator \( \varphi(p) \)?

In \( \psi(p) \) we can choose our wave function:

\[ \varphi_b(p) = \delta(p-b), \]

just in the same way as \( \psi(27) \) in the coordinate space. Here the
\( \Phi(p)'s \) are also orthonormal to one another and the function of momentum \( p \), and \( b \) is eigenvalue. Hence, by the same analogy as XI(28), we get:

\[
\Phi^*(p) = \sum_b c_b^* \delta(p-b).
\]

\[ \text{XI (34)} \]
\[
\Phi'(p) = \sum_b c_b \delta(p-b).
\]

This means that \( \Phi^*(p) \) and \( \Phi(p) \) are 0 except for \( p = b \), and creates or destroys the particle in the system where the momentum of the particle is equal to \( b \) which implies that the state is also \( b \). We can also write this as follows:

\[ \text{XI (35)} \]
\[
\Phi^*(p) = c_p^*, \quad \Phi(p) = c_p.
\]

These are the creation and destruction operators at the momentum \( p \). Generally, we can say that the operators \( \Phi(q) \) create or destroy the particle at \( q \) whatever the argument is. One can generalize the anticommutation relations of the \( C \)'s by deriving the following ones for the \( \Phi(x) \):

\[
\left[ \Phi^*(x), \Phi(z) \right]_+ = 0
\]

\[ \text{XI (56)} \]
\[
\left[ \Phi(x), \Phi(x') \right]_+ = 0
\]
\[
\left[ \Phi^*(z), \Phi(x') \right]_+ = \delta(x-x'),
\]

where the \( \Phi(x) \) is the operator which is the function of any sort of coordinate. The proof can be given by the fact that we can develop the \( \Phi(x) \) in terms of any orthonormal wave function as in XI(20) and by inserting the anticommutation relations XI(18).
In general, we can say that the $\psi$'s are the fundamental operators, whereas $C$'s are just the Fourier decomposition of $\psi$.

Now we come to the close analogy to the Bose case where the $C_k^*$ and $C_k$ are the Fourier decompositions of the operator $F(x)$. However, we should be careful to use the analogy since there are fundamental differences between the Fermi particles and the Bose particles as we know so far.

* * *

4951/E
Twelfth Lecture

(Notes collected by A. Kusumegi)

June 1958

In the previous lecture we have introduced the operators \( \psi^*(q) \) and \( \psi(q) \), which are the creation and destruction operators of a particle at the "place" \( q \), whereby \( q \) can be any kind of variable. We now plan to express the operators which correspond to physical magnitudes (energy momentum, angular momentum, etc.) in terms of those creation and destruction operators.

We first look at so-called "single particle" operators. These are operators which are a sum of operators, which operate on the variables of one particle only. An example is the kinetic energy \( T \):

\[
\text{XII}(1) \quad T = \sum_{i=1}^{N} t(q_i) \quad t(q_i) = \frac{p_i^2}{2m}
\]

Here \( t(q_i) \) is the operator of the kinetic energy of one particle. A counter example would be a potential energy which depends upon the relative distances:

\[
V = \sum_{i,k} \psi(x_i - x_k)
\]

in which the summation is carried out over two-particle operators.

The general expression for a "single particle" operator would be:

\[
\text{XII}(2) \quad F = \sum_{i=1}^{N} f(q_i)
\]

We now would like to express this operator in terms of our new operators \( \psi \) and \( \psi^* \). For this purpose we introduce the orthonormal
function system $\mathcal{F}_a(q)$ which are the eigenfunctions of the operator $f(q)$:

\[ f(q) \mathcal{F}_a(q) = a \cdot \mathcal{F}_a(q) \]
\[ f(q) \mathcal{F}_b(q) = b \cdot \mathcal{F}_b(q) \]

When we apply the operator XII(2) on the function $[a, b, c, \ldots]$, we obviously get:

\[ F \cdot [a, b, c, \ldots] = (a + b + c + \ldots) [a, b, c, \ldots] \]

This we can also express in the form:

\[ F \cdot [a, b, c, \ldots] = \sum_a a \cdot C_a^* C_a [a, b, c, \ldots] \]

where the sum goes over all eigenvalues of the operator $f(q)$. The expression $C_a^* C_a$ operating on $[a, b, c, \ldots]$ selects only those terms in the sum $\sum_a$ which belong to the states which are occupied in $[a, b, c, \ldots]$. We can write XII(4) in the form:

\[ F \cdot [a, b, c, \ldots] = \int dq \sum_a C_a^* \mathcal{F}_a(q)f(q) \sum_b C_b \mathcal{F}_b(q) [a, b, c, \ldots] \]

because of the orthonormality relations of the $\mathcal{F}$'s. This again, according to XI(19), is equal to:

\[ F \cdot [a, b, c, \ldots] = \int dq \psi^*(q)f(q) \psi(q) [a, b, c, \ldots] \]

Since any antisymmetric function $\psi$ can be written as a linear combination of different $[a, b, c, \ldots]$, we get quite generally the following expression for the operator $F$:

\[ F = \int dq \psi^*(q)f(q)\psi(q) \]

in which $F$ is expressed in terms of $\psi^*(q)$ and $\psi(q)$. One can also interpret XII(5) in a simple way: the operator $\psi^*(q)\psi(q)$ is unity at those values of $q$ where there is a particle and zero at all others. Hence the integral in XII(5) "scans" the $q$-space and interprets the values $f(q)$ over the places in the $q$-space where there is a particle.
(Remark : $q$ can be any variable; coordinate, momentum, etc.)

A special kind of operator can be constructed out of the $\psi(q)$. It is the particle density operator $\rho(q)$:

XII(6) \hspace{1cm} \rho(q) = \psi^*(q) \psi(q)

It is unity whenever there is a particle at $q$ and zero if there is not. Obviously:

\[ \int \psi^*(q) \psi(q) dq = N \]

where $N$ is the total number of particles, a relation which one can easily prove by using the expressions XII(19).

Let us now apply this new way of writing operators to the simplest many body problem, which is the one-body problem. We are now going to "derive" the Schrödinger equation for one particle with our new formalism.

We start with the "Vacuum" which we call $\mathcal{0}$. It is that state of our system for which:

XII(7) \hspace{1cm} \psi(q) [\mathcal{0}] = 0

for all values of $q$. This equation says that the vacuum is that state from which one cannot take anything away. A one particle state is created from the vacuum by:

$\psi^*(q) [\mathcal{0}]$

in which one particle is added "at $q$". The most general one particle state is:

XII(8) \hspace{1cm} \psi = \int dq \hat{f}(q) \psi^*(q) [\mathcal{0}]

where $\hat{f}(q)$ is an arbitrary function of $q$. The Hamiltonian of our system is:

\[ H = \sum_{i} h(q_{i}), \quad h(q) = t(q) + V(q) \]
where \( t(q) \) is the expression of the kinetic energy operator in terms of \( q \), and \( v(q) \) the corresponding expression of the potential energy. (We make the assumption that the potential energy is a single particle operator, which is certainly true if we are dealing with one particle only!) If \( q \) is the coordinate, we have \( t(q) = -\frac{1}{2m} \frac{\partial^2}{\partial q^2} \). We are now looking for a \( \psi \) that is an eigenfunction of the Hamiltonian:

\[
\text{XII}(9) \quad \hat{H} \psi = E \psi
\]

In our writing (see XII(5)) we have:

\[
\text{XII}(10) \quad \hat{H} = \int \psi^*(q) h(q) \psi(q) dq.
\]

Putting this and XII(8) into XII(9), we get:

\[
\text{XII}(11) \quad \hat{H} \psi = \int dq \int dq' \psi^*(q) h(q) \psi(q) \xi(q') \psi^*(q') [0].
\]

We now make use of the fact expressed by XII(7) and push \( \psi(q) \) all the way to the right so that it operates on \([0]\). Since \( \psi(q) \) does not commute with \( \psi^*(q') \) we must use the commutation relations XI(36). It commutes of course with \( \xi(q') \) which is just a plain number. We then get:

\[
\hat{H} \psi = \int dq \int dq' \psi^*(q) h(q) \xi(q') \left[ \psi^*(q') \psi(q) + \delta(q-q') \right] [0]
= \int dq \psi^*(q) h(q) \xi(q) [0] = \int dq h(q) \xi(q) \psi^*(q) [0]
\]

since the first term in the bracket gives zero according to XII(7).

Now we add the right side of XII(9) and get:

\[
\int dq h(q) \xi(q) \psi^*(q)(q) [0] = E \int dq \xi(q) \psi^*(q) [0]
\]

This relation is fulfilled if:

\[
\text{XII}(12) \quad h(q) \xi(q) = E \xi(q)
\]

which is the one-particle Schrödinger equation with \( \xi(q) \) as the Schrödinger wave function. Let us find out what the density operator XII(6) tells us in this example. We ask for the expectation value of the
density \( \rho(q_0) \) at \( q_0 \):

\[
\mathbf{XII(13)} \quad \langle \rho(q_0) \rangle_{E_x} = (\psi, \rho(q_0) \psi).
\]

Here we use the relation:

\[
\mathbf{XII(14)} \quad (Q \psi, R \psi) = (\psi, Q^+ R \psi)
\]

where \( Q^+ \) is the hermitian conjugate operator to \( Q \). \( (\psi^*(q))^+ = \psi(q) \).

Hence we get:

\[
\begin{aligned}
\langle \rho(q_0) \rangle_{E_x} &= \langle [0], \int dq \int dq' \xi^*(q) \psi(q) \psi^*(q_0) \psi(q_0) \int dq' \xi(q') \psi^*(q') \rangle [0] \\
&= \langle [0], \int dq \int dq' \{ \psi^*(q_0) \psi(q) + \delta(q_0 - q) \} \xi^*(q) \xi(q') \cdot \\
&\cdot \{ \psi^*(q') \psi(q_0) + \delta(q' - q_0) \} \rangle [0].
\end{aligned}
\]

According to relation \( \mathbf{XII(14)} \) and \( \mathbf{XII(7)} \) we find that \( \langle [0], \psi^*(q_0) q [0] \rangle = 0 \) whatever the operator \( q \) is. Hence we get:

\[
\begin{aligned}
\langle \rho(q_0) \rangle_{E_x} &= \langle [0], \int dq \int dq' \delta(q_0 - q) \delta(q' - q_0) \xi^*(q) \xi(q') [0] \rangle \\
&= \xi^*(q_0) \xi(q_0) [0], [0] = |\xi(q_0)|^2 \quad \text{since} \quad ([0], [0]) = 1.
\end{aligned}
\]

This is the expected result for a one-particle problem.
We first recall briefly what we have done and the hypotheses assumed, to emphasize the consequences of these:

**Formal Field:**

The function which describes the state of the system is antisymmetric.

The system is composed of a finite number of particles. The field is described by the operator \( \psi^j(x) \) defined as follows:

\[
\psi^j(x) = \sum_k c_k \psi^j_k(x) \quad \psi^j_k(x) = \sum_k c_k^* \phi^j_k(x) \tag{2}
\]

\[
[c_k, c^*_l] = 0 \quad [c_k, c^*_l] = \delta_{kl} \quad [\psi^j(x), \psi^j_k(x')] = \delta(x - x')
\]

\( \phi^j_k \) is a complete orthonormal system of functions so that

\[
\sum_{k, l} \phi^j_k(x) \phi^j_l(x') = \delta(x - x') \tag{3}
\]

and \( c_k \) is an operator which destroys a quantum in the state characterized by the state function \( \phi^j_k \). While \( c_k^* \) creates a quantum in the state \( \phi^j_k \), so the operator \( a c_k + b c_k^* \phi^j_k \) creates a quantum in the state defined by \( a \phi^j_k + b \phi^j_k \). In the same way we see that

\[
\psi^j_k(x) \psi^j_k(x') = \delta(x - x') \tag{4}
\]
or better, the operator $\psi^*(x)$ localizes a particle at the point $x = x^1$. Naturally the equality (4) can be established if, and only if, the sum is performed over all values of $k$, or better, if all the possible states $\psi_k$ enter into the left side of (4). This means that if one of the states $\psi_k$ is already occupied, so that from the Pauli principle it must be excluded in the sum (4) the result is no longer a function, though, always a function with a maximum at the point $x = x^1$.

In a "finite number quanta theory" such a possibility will not happen, while it certainly will happen in an "infinite number quantum theory". For instance in the hole theory of the positron, matter is always present (in the negative energy states), therefore in the sum (4) only half of the number can, at most, be present. In this case $\psi^*(x)$ is no longer creating, or localizing, a quantum at the point $x = x'$, but such a quantum can be created being distributed over a region with a maximum intensity at $x = x'$.

In a non relativistic theory of the electron, the filling up of negative energy states is disregarded. Hence a localization of a fermi quantum at the point $x = x^1$ has meaning only in the classical limit. To conclude we want to recall the expression of the hamiltonian under the hypothesis that the quanta do not interact with themselves.

$$H = \int \psi^* \mathcal{H}(x) \psi \, d^3x = \int \psi^* \frac{\partial}{\partial t} \psi \, d^3x \quad (5)$$

where $\mathcal{H}(x)$ is the hamiltonian for one particle, (see lecture 10.12) It has to be remarked that the hamiltonian is linear in (and so will be the lograngian) so that $\Pi(x) = \frac{\partial L}{\partial \dot{\psi}}$ will not contain $\psi$ (see Bose field).

Bose Field.

For what we want to clarify it is sufficient to recall the case of the uncharged boson field. It was found that the operator describing such a field could be written in the following form:
F (x) = \sum_{k} \left( \frac{c_k + c_k^*}{\sqrt{2} \omega_k} \right) f_k (x) \quad (6)
\text{\(f_k\) = real functions}
\left[ F (x), \pi (x) \right] = \delta (x - x')
\pi (x) = \frac{\partial C}{\partial F (x)}
\omega_k = \sqrt{\hbar^2 + p_k^2}
\left[ c_k, c_k^* \right] = \delta_{k, k}

\text{\(C_k\) and \(C_k^*\) are destruction and creation operators for a quantum in the state defined by \(f_k\); \(f_k, f_k', f_k'', \cdots\) is a complete orthogonal set with the condition that the \(f_k\) be eigenfunctions of a physical system with definite energy \(\omega_k\) so that to each \(f_k\) corresponds the energy of the state defined by \(f_k, \omega_k\). Let us see now if it is possible to give to \(F(x)\) the same meaning we found for \(\Psi (x)\) in the fermi field. We apply the operator \(F(x)\) upon the vacuum, then the part with \(C_k\) vanishes and we get only:}

F (x) = \sum_{k} \frac{c_k^* f_k (x)}{\sqrt{2} \omega_k} \quad (7)

\text{Proceeding now for (7), as we did for \(\Psi (x)\) we find that \(F(x)\) creates a quantum in the state characterized by}

\sum_{k} \frac{1}{\sqrt{\omega_k}} \pi (x) \frac{\partial C}{\partial F (x)} \quad (8)

\text{The sum (8) is not a \(\delta\) function because of the factors \(\frac{1}{\sqrt{\omega_k}}\); though it is a function with a singularity at \(X = X^1\) (see figure).}
however in the limiting case of non-relativistic theory

$$\omega_K = \sqrt{\mu^2 + p^2} \approx \gamma \mu$$

$F_1$ may be regarded as a $g$ function.

The Hamiltonian for such a boson field is:

$$H = \frac{1}{2} \int \left( -\nabla^2 + (\Delta F)^2 + \kappa |F|^2 \right) d^3x \quad \text{(see lecture n° G)}$$

$$L = \frac{1}{2} \int \left( +\nabla^2 + (\Delta F)^2 + \kappa |F|^2 \right) d^3x$$

Now $L$ contains $\dot{F}$ to the 2nd power so that $\frac{dL}{dF} = 2 \frac{\dot{F}}{F}$ implies $L(x) = \frac{\dot{F}}{F}$, quite differently from the fermi field.

Remark.

At this point it is useful to point out that in the quantization of the boson field we assumed as a starting point a classical field while for a fermion field we started from a classical particle. Quantizing a classical field we obtained the boson field and quantizing the particle we obtained the Schrödinger equation; after the quantization the classical field behaved more as a particle and the particle more as a field.

Some of the differences which we found have their origin from these facts and as the number of particles in the Fermi case tends to infinity (hole-theory), some of the differences between the two fields tend to disappear, so that at the end the theories are rather similar. But always when one goes to a classical limit one has to come to differences which reflect the different starting points of the two theories. Recapitulating, the differences we found were:

1). the appearance of the factor $\frac{1}{\sqrt{\omega_K}}$ in the expression of the Fourier decomposition of the $\sqrt{\omega_K}$ field.
2) the different hamiltonians assumed for the two fields,  

3) in the bose field the field operator obeys  
\[ \left[ \hat{F}(x), \hat{F}(x') \right] = \delta(x - x') , \]

while in the fermi field the field operator obeys  
\[ \left[ \hat{\Psi}(x), \hat{\Psi}^\dagger(x') \right] = \delta(x - x') . \]

Another remark which is necessary is the following: In many books the Fermi field \( \Psi(x) \) is written

\[
\Psi(x) = \sum \frac{1}{\sqrt{\omega_K}} c_K \hat{f}_K \tag{9}
\]

instead of (1) where the system of \( \hat{f}_K \) is chosen to be the wave functions of the free dirac one particle equation. The choice of the expression (9) for the \( \Psi(x) \) is possible provided that the normalization of the \( \hat{f}_K \) is chosen to be

\[
\int_{V} \hat{f}_K \hat{f}_K^\dagger d^3x = \omega_K \delta_{K_1} \tag{10}
\]

Therefore the conclusions made previously are quite valid, though the expression \( \Psi(x) \) and \( F \) look equal now. The reason some people prefer the normalization condition of the type (10) instead of the normal way:

\[
\int_{V} \hat{f}_K \hat{f}_e d^3x = \delta_{K_1} e \]
is that the latter normalization is valid only in a special Lorentz frame so that changing frames one has to rewrite the normalization condition; instead, the expression (10) is Lorentz invariant.

Of course if one starts from a theory which is Lorentz invariant and calculates expectation values of physical quantities the results are Lorentz invariant, though during the calculation one deals with expressions which are not Lorentz invariant. If however during these calculations one carries out any process of renormalization (i.e. eliminates expressions which are not observables) that part which is to be eliminated must be Lorentz invariant, otherwise the process has no physical meaning. Thus it is extremely useful in these cases to deal always with Lorentz invariant expressions.

Example

In the Schrödinger theory of a charged point particle the charge density was a number which was given by

\[ \mathcal{J} = e \int \mathcal{J}^* \]

where \( \mathcal{J} \) is the wave function of the particle properly normalized. Using the formalism previously developed we want to introduce a "charge density operator"

\[ \mathcal{J} = e \left[ \mathcal{J}^* (\mathbf{x}), \mathcal{J} (\mathbf{x}) \right] \]

(11)

It is possible to justify this choice observing that \( \mathcal{J} (\mathbf{x}) \) destroys a particle at \( \mathbf{x} \) where \( \mathcal{J}^* (\mathbf{x}) \) creates a particle at the point \( \mathbf{x} \); but \( \mathcal{J} (\mathbf{x}) \) can destroy a particle only if there is already a particle at that point so that the operator (11) gives zero when applied to a state which does not contain a particle at the point \( \mathbf{x} \). Let us now construct the operator

\[ \mathcal{J} (\mathbf{x}_1) \mathcal{J} (\mathbf{x}_2) \]

(12)
If we calculate the expectation values of this operator in a one particle system we get the probability of finding certain charge densities at \( X_1 \) and \( X_2 \); so if the particle has finite dimensions we obtain in this way the form of the particle. (However we know that in this case we must find a \( \delta \) function).

In the case of a two-particle system the final expression would be of three terms; two of them will be the shapes of the particles, the third one will give the correlation in space of the two particles.

The expectation values of the operator (12) are given by the expression - see lecture 12, (last part) -

\[
\langle 0 | \int \! d^3 x \, f (x) \, \psi (x) \, \psi^* (x_1) \, \psi (y) \, \psi^* (x_2) \! \int \! d^3 y \, f (y) \, \psi^* (y) | 0 \rangle
\]  

in the same lecture it is shown that

\[
\psi = \int \! d^3 x \, f (x) \, \psi^* (x) [0]
\]

is the most general form of the one particle state

\[
f (x) = \text{Schrödinger wave function}
\]

\[
[0] = \text{state which represents the vacuum characterized by:}
\]

\[
[0] \psi = 0 \quad [0] \psi^* = 0
\]

In the expression (13) we can first, through the equality

\[
\psi (x) \psi^* (x_1) + \psi^*(x_1) \psi (x) = \delta (x - x_1)
\]

\[
\psi (x_2) \psi^* (y) + \psi^*(y) \psi (x_2) = \delta (y - x_2)
\]

bring \( \psi (x_2) \) to the right and \( \psi^* (x_1) \) to the left; then using (14) - the same technique was used in the 12th lecture - we obtain:

\[
\langle 0 | \int \! d^3 x \, f (x) \, \delta (x - x_1) \, \psi (x_1) \, \psi^* (x_2) \! \int \! d^3 y \, f (y) \, \delta (y - x_2) | 0 \rangle
\]
Repeating the operation just applied we arrive at

\[ \langle 0 | \hat{f}(x_1) \Psi(x_1) \Psi^*(x_2) \hat{f}(x_2) | 0 \rangle = \]

\[ = \langle 0 | \hat{f}(x_1) \delta(x_1-x_2) \hat{f}(x_2) | 0 \rangle = \left| \hat{f}(x_1) \right|^2 \delta(x_1-x_2) \]

As we expected, we obtain a probability different from zero only if \( x_1 = x_2 \) (point particle).

We will see in one of the next lectures that in an infinite number of quanta theory (for instance in the hole theory of electrons), because of the vacuum (whose behaviour is influenced by the presence of the particle), the expectation value of the expression (12) - in the case of one particle in the vacuum - is no longer a \( \delta \) function. However, it will be a function with a maximum for \( x_1 = x_2 \), having there a singularity.
We found that the Dirac equation for a single electron could be written in the Maxwell form as
\[
(\vec{\sigma} \cdot \vec{E}) \psi + m \psi = E \psi \tag{1}
\]
\[
(\vec{\sigma} \cdot \vec{p}) \psi - m \varphi = E \varphi. \tag{2}
\]

The two component spinors $\psi$ and $\varphi$ together form the four-component spinor
\[
\Psi = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \tag{3}
\]

For a given momentum eigenvalue $p$ there are in general four independent solutions of the type
\[
\Psi = u(p,E) \exp (ipx - iEt) \tag{4}
\]
if we choose the $z$-axis in the direction of motion of the particle. The various possibilities are two different polarization states for each of the two possible energy eigenvalues given by
\[
E = \pm \sqrt{p^2 + m^2} = \pm \frac{1}{2} \sqrt{M^2}
\]
Choosing the two directions of polarization to be either parallel ('UP' or $\uparrow$) or antiparallel ('DOWN' or $\downarrow$) to the positive $z$-axis we obtain the following table for the four four-component spinors $u(p,E)$

<table>
<thead>
<tr>
<th>$u \times \frac{W+m}{2W}$</th>
<th>1</th>
<th>0</th>
<th>$-p/(W+m)$</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p/(W+m)$</td>
<td>o</td>
<td>1</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>o</td>
<td>-$p/(W+m)$</td>
<td>o</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spin</th>
<th>Energy</th>
<th>$E=W$</th>
<th>$E=-W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\uparrow$</td>
<td>$E=W$</td>
<td>$E=W$</td>
<td>$E=-W$</td>
</tr>
</tbody>
</table>

Table 1.
The factor \( \left( \frac{W + m}{2W} \right)^{\frac{1}{2}} \) is a normalization factor such that
\[
\sum_i |\psi_i(p, E)|^2 = 1
\]
In general the two-component spinors \( \psi \) and \( \varphi \) define polarization directions lying symmetrically about the z-axis, but with our special choice of axis they are either both parallel or both anti-parallel to the z-axis.

That we may have solutions with negative energy eigenvalues is a striking fact of the Dirac theory. We note that even in the classical theory of relativity negative energies could be deduced from expression (5), but there they could not be related to any physically observable states, since \( E \) had to vary continuously and the gap between \( E \geq m \) and \( E \leq -m \) could not be bridged. In quantum theory such transitions can take place, and the negative energy states arising from the Dirac theory can therefore no longer be ignored.

In order to see how these negative energy solutions should be interpreted we introduce an electromagnetic field given by its potentials \( \vec{A} \) and \( \varphi \). It can be shown that the effect of this field on the Dirac equation is equivalent to replacing
\[
\vec{p} \quad \text{by} \quad \vec{p} + e\vec{A}
\]
and
\[
E \quad \text{by} \quad E + e\varphi,
\]
where the upper signs are valid for a positively charged particle and the lower signs for a negatively charged one, \( e \) is the magnitude of the charge. No proof of this substitution is offered here, but its justification can be found in any Quantum-mechanics course.

We call the solutions of the Dirac equations in the presence of an EM field \( \Phi \) for the positive particle and \( \Psi \) for the negative one. The solutions \( \Phi \) and \( \Psi \) will in general be different, but one can show\(^*)\) that for any field \( \vec{A}, \varphi \) there exists between them the relation
\[
\Phi_E = C \Psi_E^*.
\]
\(^*)\) See notes on Dirac equation.

5237/e/p
Starring in (6) indicates the complex conjugate, not the Hermitian conjugate. C is the four by four matrix

\[
C = \begin{pmatrix}
0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 0
\end{pmatrix}
\]  

Thus the positive energy solutions of one charge are related to the negative energy solutions of the opposite charge. In fact, one only has to take the conjugate complex of the negative energy solutions (of course, in order to change the sign of the energy from negative to positive in the factor \( e^{-iE t} \)) and to reverse the order of the components \( \phi_1(p, E) \) apart from a few minus signs. The latter operation is what the matrix C achieves. \( \Phi_{-E} \) is the charge conjugate of \( \Psi_E \).

As an illustration we consider the Dirac solutions for the hydrogen atom. For positive energy of the electron we got the Sommerfeld spectrum of the relativistic electron: the well-known discrete energy levels below \( E = m \) and a continuum above (see figure). The levels with negative energies can be guessed at by remembering that they must (apart from the sign) be identical with the levels of a positron in the field of the proton.

This would give rise to a continuum of eigenvalues from \( E = m \) up to infinity, since there are no bound states with discrete levels in such a system. Hence the negative energy levels of an electron start from \( E = -m \) as a continuum until \(-\infty\).

As another illustration, let us look at the relation (6) for no field, \( \vec{r} = \phi = 0 \); that is for a free particle. In this case (since the charge is then irrelevant) the charge conjugation (6) must lead to another state of the same particle. Let us consider one of
the solutions with negative energy \(-E\), say the one with spin up:

\[
\hat{\Psi}_E^r = \begin{pmatrix}
-p/(\bar{W}+m) \\
0 \\
1 \\
0
\end{pmatrix}
\begin{pmatrix}
-W+m \\
2W
\end{pmatrix}^{\frac{1}{2}} e^{i\frac{\vec{p} \cdot \vec{x}}{2} + iWt}
\]

Let us now apply to it the operation (6):

\[
\hat{\Phi}_E = C\hat{\Psi}_E^r = \begin{pmatrix}
0 \\
1 \\
0 \\
+p/(\bar{W}+m)
\end{pmatrix}
\begin{pmatrix}
-W+m \\
2W
\end{pmatrix}^{\frac{1}{2}} e^{-i\vec{p} \cdot \vec{x} - iWt}
\]

We see then that, as expected, \(\hat{\Phi}_E\) is a positive energy solution. We also note that it is the positive energy solution for opposite momentum \(-p\) and opposite spin (down).

This also leads to the statement that the negative energy states must be read in time reversal in order to be physically significant, time reversal changes the sign of energy (factor \(e^{-iEt}\)); but it also changes the direction of momentum and spin. Hence, the negative energy states are physically quite meaningful. One must read them the correct way, that is by means of operation (6) which corresponds to time reversal. Then, they give us the positive states of the particle with opposite charge.

There are two ways of dealing with the negative energy states:

A). Dirac Interpretation, hole-theory.

Dirac assumes that the negative energy states of the electron are physically meaningful but are normally occupied by the "Vacuum electrons." The vacuum is therefore the state in which all negative energy states are filled. If one of the states is unfilled, we have a hole, which will act like a positron. A "lacking" state of energy \(-E\), momentum \(p\) and spin \(\sigma\) is, of course, a positron state of energy \(+E\), momentum \(-p\), spin \(-\sigma\). This is just like the transformation (6). The element of time reversal also enters into it,
since a transition of a vacuum electron from a state $\alpha$ to $\beta$ ( $\beta$ being unfilled) corresponds to the **reversed** transition of a positron, from $\beta$ to $\alpha$. The difficulties of the Dirac representation come from the infinite charge - and energy-density of the vacuum. Actually, however, a uniform charge density produces no forces and is therefore unobservable in principle. An infinite uniform energy density is also a type of difficulty which we have previously decided to ignore e.g. the infinite zero point energy encountered in the quantization of the Boson field. Another way out of the infinitesimal charge density is to consider the vacuum as a linear combination of two worlds: in one the negative energy states are occupied by positrons, in the other by negatrons.

It is worth noticing that the ripples in space or time of the charge density due to fluctuations or to polarization effects of external fields are observable.

**B). Feynman Interpretation**

Feynman does not fill the negative states with vacuum electrons but **presumes to read** the negative energy states in the opposite time direction. This makes them positive energy states of the other charge. Certain formal complications occur since it is somewhat difficult to use one wave equation and to read the time in the opposite direction for half of the states. It can be done, however.

We will not use Feynman's method in this course, but we will give one simple example. Fig. 2a shows a time record of the following process: an electron is at the times $t < 0$ in a state $\alpha$ with positive $E$. At $t = 0$ it emits a light quantum (or two) and jumps into a state $\beta$ with negative energy $-E'$. Feynman tells us to read the time backwards when it is in $\beta$. We then get Fig. 2b in which we see for $t < 0$ an electron in the state $\alpha$, and a positron in the state $\beta^*$ and energy $+E'$. At $t = 0$ light quanta are omitted and for $t > 0$ there is nothing but the light quanta.
Hence we get a perfect picture of annihilation.

Fig. 2a

Fig. 2b.
Fifteenth lecture

(Notes collected by E.G. Michaelis and H. Øverås)

We shall now consider more closely the fluctuations in vacuum, and we shall stick to the Dirac method.

We have to do with a system of infinitely many Fermi particles, and the problem is not fundamentally difficult, but it is difficult in detail.

According to the degree of approximation the problem may be formulated in three steps:
2. Outside field considered.
   What is the effect of an outside field, say from proton, on the vacuum, and what is the influence on the electrons of this perturbed vacuum?
3. All electromagnetic effects taken into account. Not only outside fields, but also the fields created by the electrons (normal and vacuum) taken into account. The problem is very difficult, and it must be attacked with high-power quantum electrodynamics. The forces between negative energy electrons are certainly very strong, but they still lead to rather small perturbations of the system, because most transitions are prevented by the fact that neighbouring states are generally occupied. It is the same reason that makes the independent nuclear model for the nucleus work so well.

With the approximation 1 and 2 we shall consider the problem of infinite self energy and infinite self charge.
Self Energy

We shall limit ourselves to the electrostatic self energy, because the problem is then simpler, but still contains the essential features.

The problem is an old one, and the situation is rather improved when taking vacuum fluctuations into consideration.

Classically the electrostatic self energy is \( \propto \frac{e^2}{\alpha} \), where \( \alpha \to \infty \) leads to infinite energy. In order to avoid that, one assumed that either the electron had an extension of order \( \alpha \), or that the electrodynamics had to be modified at such distances, \( \alpha \) being determined by the condition

\[
\frac{e^2}{\alpha} \leq mc^2 \quad \text{or} \quad \alpha > \frac{e^2}{mc^2}.
\]

This would mean that electrodynamics does no longer hold for momenta higher than \( p_\alpha \) given by:

\[
\lambda := \frac{\hbar}{p_\alpha} \quad \text{or} \quad p_\alpha \approx \frac{127mc}{\lambda} \approx 70 \text{ MeV/c}.
\]

This is rather low value and, fortunately, that classical conclusion will turn out invalid.

We shall now calculate the self energy of one electron, neglecting the influence of the vacuum electrons, in a formal quantum mechanical way. The state of 1 electron present, say in the ground state we call \( | 1_0 \rangle \) and the wavefunction for the electrons \( \varphi_n(\mathbf{x}) \).

5239/e/p
The electrostatic self energy is then

$$E_{\text{self}} = \frac{1}{2} \int \frac{\tilde{G}(\vec{x}_1 - \vec{x}_2)}{|\vec{x}_1 - \vec{x}_2|} \, d^3x_1 \, d^3x_2$$

(1)

The function $\tilde{G}(\vec{x}_1 - \vec{x}_2)$ expressing the correlated probability of finding simultaneously one electron at $\vec{x}_1$ and one at $\vec{x}_2$, is given as the expectation value of the product of the two density operators $\varrho(\vec{x}_1)$ and $\varrho(\vec{x}_2)$:

$$\tilde{G}(\vec{x}_1 - \vec{x}_2) = \langle 1_0 | \varrho(\vec{x}_1) \cdot \varrho(\vec{x}_2) | 1_0 \rangle.$$ 

The density operator is given by $\varrho(\vec{x}) = -e \left\{ \Psi^\dagger(\vec{x}) \Psi(\vec{x}) \right\}$, where $\left\{ \Psi^\dagger(\vec{x}) \Psi(\vec{x}) \right\}$ is the spinor product

$$\sum_\alpha \Psi^\dagger_\alpha(\vec{x}) \Psi_\alpha(\vec{x})$$

denoting the four spinor components. We have

$$\Psi(\vec{x}) = \sum_c c_c \varphi_c(\vec{x}).$$

So

$$\tilde{G} = e^2 \langle 1_0 | \sum_{i,k,l,m} c_i^* \{ \varphi_k^\dagger(\vec{x}_1) \varphi_l(\vec{x}_1) \} \{ c_i^\dagger \varphi_k^\dagger(\vec{x}_2) \varphi_l(\vec{x}_2) \}| 1_0 \rangle.$$ 

The action of the destruction and creation operators $C$ and $C^*$ on the state $| 1_0 \rangle$ may be illustrated by the adjoining diagram:

```
          l = k
            ^
            |
            v
          i = m = o
```

We only get such terms in the sum where

$$i = m = o, \quad k = l.$$ 

5239/e/p
Then

\[ G = e^2 \sum_i \{ \Phi_i^*(x_i) \Phi_i(x_i) \} \{ \Phi_i^*(x_2) \Phi_i(x_2) \} . \]

Now using the completeness relation

\[ \sum \Phi_i^{(\alpha)}(x_i) \Phi_i^{(\beta)}(x_2) = \delta(x_i - x_2) \delta_{\alpha \beta} \]

we get

\[ G = e^2 \delta(x_i - x_2) \mid \Phi_i(x_i) \mid^2 . \]

Inserting this expression in the self energy integral:

\[ E_{\text{self}} = -\frac{e^2}{2} \int \frac{\delta(x_i - x_2)}{|x_i - x_2|} d^3 x_i d^3 x_2 \]

\[ = -\frac{e^2}{2} \int \frac{\delta(x_i - x_2)}{|x_i - x_2|} d^3 (x_i - x_2) . \]

This integral diverges. It is identical with the electrostatic energy of a charge distribution \( \delta(x) \). If we replace the \( \delta \) function by a finite distribution with radius \( \alpha \), we get \( E_{\text{self}} \propto \frac{e^2}{\alpha} \). In the limit \( \alpha \to 0 \) we obtain an infinity just as in the classical theory.

This is to be expected since the one particle quantum mechanics assumes the particle to be specially confined to a "point"
Only its probability distribution is smeared out. Some people make the mistake of interpreting the probability distribution (say in the ground state $\varphi_0$) as an actual charge distribution. Then the self energy would be

$$E_{\text{self}} = e^2 \int \frac{\varphi_0(x_1)^2 \varphi_0(x_2)^2}{|x_1 - x_2|} \, dx_1 \, dx_2$$

This expression would depend upon the form of the wavefunction, and a different self energy would have to be added to the various energy levels. This would change completely the spectrum, say, of the hydrogen atom.

For the hydrogen ground state, e.g., $E'_{\text{self}}$ calculated in this way is $\frac{e^2}{\hbar \text{Bohr}}$, which is finite, but not right.

We therefore go back to the concept of a well confined electron, and with the approximation 2), taking into account the filled negative energy states, we shall calculate again the self energy of the electron, showing that the situation is greatly improved, in fact that the critical length $a$ may be reduced to

$$\frac{\hbar}{mc} e^{-1/37}$$

The presence of an electron modifies the vacuum, and this effect cannot be separated from the properties of the electron itself. We therefore have to add to the electron, also the effects of the modification of the vacuum. The correlated charge distribution is then the difference between $G(x_1 - x_2)$ when our given electron is added to the vacuum electrons and the value of $G(x_1 - x_2)$ for the vacuum electrons by themselves:

$$\tilde{G} = \langle 1_0 \text{vac} | G | 1_0 \text{vac} \rangle - \langle \text{vac} | G | \text{vac} \rangle.$$  \hspace{1cm} (2)
We again write
\[
\langle A|G_A|A \rangle = \langle A| \sum_{i \in \mathcal{E} \cap m} \left\{ c_i^* \varphi_e^* (x_i) c_i \varphi_e (x_i) \right\}_i \left\{ c_e^* \varphi_e^* (x_e) c_e \varphi_e (x_e) \right\}_e \rangle_A
\]
where the state A is either " \( |\psi \rangle \) vac" or " vac \( |\psi \rangle \)"

There are again restrictions in the values of \( i, k, \mathcal{E}, m \), since the four successive operators must re-stitute the original state:
\[
c_i^* c_k^* c_e^* c_m = 1 \quad \text{this is the case only if } i = m, k = q, i \text{ occupied, } k \text{ unoccupied, or } i = k, e = m, i \text{ occupied, } k \text{ occupied.}
\]

if one goes through this operation one finds that (2) becomes
\[
\bar{G} = e^2 \left( \sum_\mathcal{E}^+ - \sum_\mathcal{E}^- \right) \left\{ \varphi_e^* (x_1) \varphi_e (x_1) \right\}_e \left\{ \varphi_e^* (x_2) \varphi_e (x_2) \right\}_e
\]
\[
+ e^2 \sum_\mathcal{E}^- \left\{ \varphi_e^* (x_1) \varphi_e (x_1) \right\}_e \left\{ \varphi_e^* (x_2) \varphi_e (x_2) \right\}_e
\]
\[
\text{(3)}
\]
where the symbol \( \sum^\pm \) means summation over states of positive or negative energy respectively. The second term must be dropped, because it leads to the electrostatic energy of the charge distribution \( | \varphi_e (x_1) |^2 \) imbedded in the charge distribution \( \sum_\mathcal{E}^- | \varphi_e (x_2) |^2 \)

of the vacuum electrons. This effect is supposed to be omitted.

We then get
\[
\bar{e} = e^2 \left( \sum_\mathcal{E}^+ - \sum_\mathcal{E}^- \right) \left\{ \varphi_e^* (x_1) \varphi_e (x_1) \right\}_e \left\{ \varphi_e^* (x_2) \varphi_e (x_2) \right\}_e
\]

In contrast to the previous
\[
\left( \sum_\mathcal{E}^+ - \sum_\mathcal{E}^- \right) \varphi_e^* (x_1) \varphi_e (x_1)
\]
\[
\text{(4)}
\]
is no longer a \( \delta \)-function. It would be one if we had

5239/e/p
\[ (\Sigma^+ + \Sigma^-) \quad \text{. It can be easily calculated by using the expressions for} \quad \Psi_e \quad \text{given in lecture 14. Actually} \]

\[ \Psi_\epsilon - (\hat{\kappa})\Psi_\epsilon + (\hat{\kappa}) - \Psi_\epsilon - (\hat{\kappa})\Psi_\epsilon + (\hat{\kappa}) = \left[ \frac{1 - \frac{|p|^2}{(w + m)^2}}{2w} \delta_{\epsilon(p)} \right] \Psi_\epsilon - (\hat{\kappa}) \]

\[ = \frac{m}{w} \delta_{\epsilon(p)} \Psi_\epsilon - (\hat{\kappa}) \quad \text{Hence we get for (4)} \]

\[ \sum_{\Psi} \frac{m}{\sqrt{p^2 + m^2}} e^{i\hat{p} \cdot (\hat{x}_1 - \hat{x}_2)} \]

in which the sum goes over all momentum states in our volume.

This sum is not a \( \delta \) -function but a Hankel function

\[ \tilde{G}_\epsilon = D = \frac{1}{4\pi \lambda_\epsilon} \frac{H^{(1)}(\epsilon \lambda_\epsilon)}{\lambda_\epsilon} = \text{Compton wavelength} \]

where \( r = |\hat{x}_1 - \hat{x}_2| \).

\( D(\hat{x}_1 - \hat{x}_2) \) has got the following asymptotic behaviour

\[ r \ll \lambda_\epsilon \quad | \quad r \gg \lambda_\epsilon \]

\[ D \sim \frac{1}{4\pi} \frac{1}{r^2 \lambda_\epsilon} \quad | \quad D \sim e^{-r/\lambda_\epsilon} \]

From this we got for the self energy according to (1):

\[ E_{\text{self}} = \lim_{a \to 0} \int_0^\infty e^2 D(r) \frac{d^3x}{r} \approx \lim_{a \to 0} \frac{e^2}{\lambda_\epsilon} \int_0^\infty \frac{dr}{r} = \lim_{a \to 0} \frac{e^2 \log \lambda_\epsilon}{a} \]
This again diverges, but only logarithmically. If we want to have a finite self energy we must introduce a finite cut off radius \( a \). We evidently would claim that the electrostatic self energy should be less than the total mass-energy: \( E_{\text{self}} < mc^2 \)

Hence we get:

\[
\frac{e^2}{x_c} \log \frac{x_c}{a} < mc^2
\]

since \( \left( \frac{e^2}{x_c} \right) = \frac{1}{137} mc^2 \), we obtain:

\[
\frac{1}{137} mc^2 \log \left( \frac{x_c}{a} \right) < mc^2
\]

\[
\log \frac{x_c}{a} < 137
\]

\[
a > x_c e^{-137}
\]

This is a very small distance and it is not surprising if electrodynamics break down at such a distance.

There is a simple interpretation of the fact that in the positron theory the correlation function \( \mathcal{D}(\vec{x}_1 - \vec{x}_2) \) is not a \( \delta \) -function. It comes from the Pauli-principle. If an electron is put into the empty space, the vacuum electrons must recede from the electron since two electrons (of the same spin) must be apart from each other by about \( \hbar/p \) where \( p \) is their relative momentum. Hence, there is a hole in the vacuum electron distribution at the place of the electron which compensates its charge. But the receding vacuum electrons crowd around it and produce the distribution \( \mathcal{D}(x) \) which we have calculated.
Sixteenth lecture

(Notes collected by F.J.M. Farley)

The last time we considered the static self-energy of the electron and calculated the effect of vacuum polarization. This gave the distribution function \(D(x_1, x_2)\) indicated in figure 1.

![Figure 1.](image)

The singularity is less strong than the delta function, and therefore we obtain a smaller self-energy:

\[
E = \int \frac{\nu \left( x_1 \right) \nu' \left( x_2 \right)}{x_1 x_2} \, c(x_1) \, dx_1 \, dx_2 \quad -(1)
\]

and only goes logarithmically to infinity as the cut-off distance \(a\) goes to zero:

\[
E = \frac{\epsilon^2}{\hbar c} m_c^2 \log\left(\frac{l}{m_c^2 a}\right) \quad -(2)
\]

There are two remarks to make about this:

**First remark:** This is only the static part of self-energy; we have also energy due to currents (i.e. spin effects). This, however, leads to self-energy of the same order of magnitude. The calculation involves coupling between light quanta and the Coulomb fields; this is not difficult but is not covered in the present course of lectures.

**Second remark:** If the electron is spread out, should the whole physics of the electron be changed? For example, is this involved in Hofstadter's experiments on high energy electron
scattering by the proton, which involve distances of order 1 fermi or less? The width of the D-function is of the order Compton wavelength of the electron, i.e. 400 fermi. How does it come about that Hofstadter's experiments can be dealt with on a point charge basis?

**Answer:** Consider the following Gedanken experiment. We have a vacuum: suddenly at time zero we put an electron in the vacuum. It will then take some time for the stationary distribution \( \psi(x_1, x_2) \) to develop; especially is this true for the wings of the distribution which involve wave functions of low relative momenta. The time for the distribution to develop is of the order \( \hbar/mc^2 \). Therefore at the first moment we have only the bare electron and then later the \( \psi \)-distribution builds up.

In the Hofstadter experiment we have very close encounters between electron and proton and these go very fast. When the impact parameter is only a few fermis, the collision time, especially when we remember the Lorentz contraction, is much less than \( \hbar/mc^2 \). Therefore the D-function does not have time to take part in the scattering. The correlation between the incident electron and the vacuum fluctuations does not adjust to the new state of motion after scattering until later. In other words, for interaction times very short with respect to \( \hbar/mc^2 \) the correlation with the vacuum electrons is not relevant. The correlation is re-established once the electron gets onto its new path. However, the closer in part of the D-function, which involves short times, can still play a role and this gives rise to some radiative corrections of order a few percent.

We note here that the radiative corrections for electron scattering are a maximum for energies \( \sim mc^2 \). For very slow collisions the impact parameter is much greater than \( \hbar/mc \), (i.e. than the width of the D-function) and there is no
effect. On the other hand, for very close collisions, i.e. at high energy, the collision time is very short and again there is no effect. The effect is therefore a maximum in the intermediate region which one would expect to be for energies \( \sim mc^2 \).

There are also corrections to the scattering from spin and magnetic energies which we have not considered herc. When we put in all these corrections for the Hofstadter experiment, we get a corrected Mott-scattering formula which applies to the scattering of a point electron by a point proton.

The experiments show a deviation from this formula. This means either (a) that the proton is not a point, or (b) that the electron is not a point. At present (a) seems more plausible since we know that the meson cloud around the proton should spread its charge. (b) is less plausible since we saw that the necessary cut off of electrodynamics has a very much smaller lower limit.

Third remark: the Lamb Shift: The Lamb Shift is the self-energy of the electron calculated in the Coulomb field of the proton, minus the self-energy of a free electron. Call \( D \) the correlation function for free electrons, and \( D_c \) the same function in the Coulomb-field. We then have:

\[
E_{\text{free self}} = \frac{1}{2} \int \frac{\left( \frac{1}{x_1 - x_2} \right)}{|x_1 - x_2|} dx_1 dx_2 \quad \text{and diverges logarithmically}
\]

but

\[
E_{\text{Coulomb self}} - E_{\text{free self}} = \frac{1}{2} \int \frac{1}{|x_1 - x_2|} dx_1 dx_2 \quad \text{and is finite}
\]

because the divergences are due to wave functions of high momenta and these are the same with and without the Coulomb field and so give zero contribution to the Lamb shift.
This is in a nutshell the calculation of the Lamb shift, but we need also to take account of the spin self energies in a complete calculation.

Fourth remark: We have now studied the correlation between our electron and the vacuum electrons and got the sort of effect indicated in Figure 1. This itself has nothing to do with the charge of the electron. It would be equally true for neutral particles. But for a charged particle, Figure 1 implies an effect on the self-energy, equation (2). This leads to the "renormalization of the mass" which means a cut-off at distance a. Here I would like to quote from a recent article by Sidney Drell, The present situation in Electro-Dynamics, .......... "Renormalization has established a peaceful coexistence between divergences and the results that you can calculate." The discrepancy is still there, but you can live with it.

Mass Renormalization for Pions: In calculating the correlation between a boson and the corresponding vacuum pair fluctuations we find that the vacuum particles are attracted to the particle instead of repelled from it as in the case of the electron. This makes the self-energy problem worse instead of better. The cut-off momentum is of order \((137)^{\frac{1}{2}} \cdot c\) and corresponds to energies of the same order as the proton mass, which is an awkward order to magnitude.

Hence, peaceful coexistence does not apply for \(\pi\) mesons.
Polarization of Vacuum, Renormalization of Charge.

We can now consider the vacuum with an outside electric source in it, (e.g. a proton). The electric charge gives a polarization of the vacuum, that is, the charge distribution in the vacuum is changed by the presence of the proton; the vacuum behaves like a dielectric, but it is more complicated. The result is that the outside charge will change its character. The effect arises because the filled negative energy states of the vacuum are now the negative energy states in the field of the proton, instead of the negative energy states of the free electron.

The difference is small because so many states are filled and the effects tend to cancel. This is the stiffness of a Fermi gas which we encounter also in the theory of nuclear matter. (If all states were filled, then external forces would not change anything at all).

One might say that because the negative energy states are the same as positron states of positive energy, and because the positrons are repelled by the proton, the vacuum electrons will always stay away from the proton. But this is not really true.
Consider a region of repulsive potential $V$ which has a maximum value $V_m$ as indicated in figure 1. What effect will this have on the electron density? If $E$ is less than $V_m$, the electron is repelled from the potential, and the electron density decreases. But if $E$ is greater than $V_m$, the electron density rises in the region of the potential due to the slowing down of the electrons. Summing over all states leads to exact cancellation. That is, on the average over all states the particle density in the repulsive region is unchanged. It will therefore, be unchanged if all states are filled. However, in our case a little effect is left due to induced pair creation, (spilling over to unfilled positive energy states), and this is the observed effect.

The proton is surrounded by a little negative cloud, because the pair electron will tend to be near the proton and the positron will tend to be farther away.

The effect is small, but infinite! That is, it diverges to infinity at small distances.

To calculate the effect we make a Fourier analysis of the $Z/r$ potential and consider first in 1-dimensional approximation the component

$$V(x) = V_o \exp (ikx) + \exp (-ikx) = 2V_o \cos kx$$

What charge density, $\rho'$, does this static field induce?
\[ \rho' = e \sum_{\ell} \left[ \psi_{\ell}^*(x) \psi_{\ell}^*(x) - \psi_{\ell o}^*(x) \psi_{\ell o}^*(x) \right] \] 

Where \( \psi_{\ell o}(x) \) = initial wave function of the \( i \) th electron and \( \psi_{\ell}(x) \) = the same wave function in the presence of the external field. The sum is taken over all filled, i.e. negative energy, states.

If we write \( \psi_{\ell}(x) = \psi_{\ell o}(x) + \psi_{1\ell}(x) \) we have

\[ \rho' = e \sum_{\ell} \left[ \psi_{\ell o}^*(x) \psi_{\ell o}^*(x) + \psi_{1\ell}^*(x) \psi_{\ell o}^*(x) \right] \] 

Perturbation theory tells us that the difference in the wave functions due to the external field is

\[ \psi_{\ell}(x) = \sum_{\ell} \frac{(e H' / \ell)}{\varepsilon_{\ell} - \varepsilon_{\ell o}} \psi_{\ell o}(x) \]

summed over all \( \ell \) not equal to \( \ell \) where

\[ H' = eV(x) = eV_0 \cos kx \]

We need only sum over the positive energy states (i.e. unfilled states), because for two filled states we have two equal and opposite contributions which cancel, due to opposite signs in the denominator.

Therefore

\[ \psi_{1\ell}(x) = -\sum_{\ell} \frac{(e V / \ell)}{\varepsilon_{\ell} + \varepsilon_{\ell o}} \psi_{\ell o}(x) \] 

as \( E_{\ell} \) is negative and \( E_{\ell o} \) is positive, and where the sum is to be taken only over positive energy states \( \ell \).
Now,

\[ \langle \ell | e V | \alpha \rangle = \epsilon \int \varphi_{\ell \alpha}^{(\alpha)}(x) V(x) \varphi_{\ell \alpha}^{(\mu)}(x) dx \]  

-(9)

and \( \varphi_{\ell \alpha} \) etc., are plane waves of form \( \exp(p_{\ell} \cdot x) \)

The integral is non-zero only if \( p_{\ell} = p_{\prime} + k \) where \( p_{\prime} \) is the momentum of the \( \ell \)-state, etc. In effect, we only have to consider matrix elements for states whose momenta differ by \( k \).

We now have to remember that each wave function involves a spinor, so we write

\[ \varphi_{\ell \alpha} = v(p_{\ell}) \cdot \exp(ip_{\ell} \cdot x) \]  

-(10)

\( v(p_{\ell}) \) being the spinor.

In evaluating a matrix element we therefore have to calculate the product of two spinors, each spinor being one column of the following matrix:

\[
\begin{array}{cccc}
1 & 0 & -p_{\ell}/W & 0 \\
0 & 1 & 0 & p_{\ell}/W \\
p_{\ell}/W & 0 & 1 & 0 \\
0 & p_{\ell}/W & 0 & 1 \\
\end{array}
\]

energy | + | + | - | - |

spin | ↑ | ↓ | ↑ | ↓ |

where \( W = (p^2 + m^2)^{1/2} + m \) (*)

(*) Attention: In the previous lectures we have had a different meaning of \( W \), namely \( W = (p^2 + m^2)^{1/2} \).
We only have to consider the products of a positive energy state and a negative energy state, and the only combinations not equal to zero are indicated by the arrows above the matrix. We see here that the spin cannot change for transitions induced by a static potential. The left-hand pair of spinors gives the product

\[ \psi^*(p^1) \psi(p_\ell) = -p^1/W^1 + p_\ell / W_\ell \]  

-(11)

Substituting (11) in (9)

\[ (\ell/eV/1) = eV_0 \left[ p_\ell / W_\ell - p^1/W^1 \right] \]

as the integral over the product of space wave functions is 1, (we do not worry about signs because we know the sign of the result)

Substituting now in (8) we obtain

\[ \Phi_{1\nu}^{(n)} = eV_0 \sum_{\ell} \frac{p_\ell/W_\ell - p^1/W^1}{|E_\ell| + |E_\ell|} \Phi_{1\ell}^{(n)} \]

-(12)

We now remember that the state \( i \) is a negative energy state and \( \ell \) is a positive one. They have the same spin but differ in momentum by \( k: p_\ell = p^1 \pm k \)

Hence the sum over \( \ell \) contains really only two terms

\[ \Phi_{1\nu}^{(n)} = eV_0 \left[ \frac{p_\ell/W_\ell - p^1/W^1}{|E_\ell| + |E_\ell|} \Phi_{1\ell}^{(n)} + \frac{p_\ell'/W_\ell' - p^1/W^1}{|E_\ell'| + |E_\ell'|} \Phi_{1\ell'}^{(n)} \right] \]

-(12a)
Where $\ell$ is the positive energy state $p_1 + k$ and $\ell'$ the one with $p_1 - k$. Let us now call
\[ f_{\ell}(k) = \frac{\epsilon_\ell/k \epsilon_\ell/k}{|\epsilon_\ell| + |\epsilon_\ell|} \]

\[ f_{\ell'}(-k) \] is the corresponding function for the state $\ell'$. Both $\ell$ and $\ell'$ states are of positive energy, so that
\[ \Phi_{0\ell} = v_+ (p_1 + k) \exp \left\{ + (p_1 + k) \times \frac{3}{2} \right\} \]
\[ \Phi_{0\ell'} = v_+ (p_1 - k) \exp \left\{ + (p_1 - k) \times \frac{3}{2} \right\} \]

Where $v_+(p)$ is the spinor belonging to a positive energy. Now we get from (12a)
\[ \Phi_{1\ell}(x) = e V_0 \left[ f_{1\ell}(k) v_+ (p_1 + k) \exp \left\{ + (p_1 + k) \times \frac{3}{2} \right\} \right. \]
\[ + \left. f_{1\ell'}(-k) v_+ (p_1 - k) \exp \left\{ + (p_1 - k) \times \frac{3}{2} \right\} \right] \]

We now put this into (6),
\[ \rho' = e \sum \left( \Phi^*_{0\ell} \Phi_{1\ell} + \Phi^*_{1\ell} \Phi_{0\ell} \right) \]
with
\[ \Phi_{0\ell} = v_+ (p_1) \exp \left\{ + p_1 \times \frac{3}{2} \right\} \]
and we get
\[ \rho' = 2 e^2 V_0 \left[ \sum f_{1\ell}(k) v_+ (p_1) v_+ (p_1 + k) \exp \left\{ + k \times \frac{3}{2} \right\} \right. \]
\[ + \left. \sum f_{1\ell'}(-k) v_+ (p_1) v_+ (p_1 - k) \exp \left\{ - k \times \frac{3}{2} \right\} \right] \]

(The second term of (6) turns out to be the complex conjugate of the first which is written above. Later we shall put $\exp(ikx) + \exp(-ikx) = 2 \cos kx$, anticipating this we put 2 outside the bracket).
Again we can write
\[ g(\pm k) \equiv \langle \tilde{\psi}(p_x) \bar{\psi}(p_x \pm k) = \frac{p^2}{\hbar c} - \frac{p_\xi}{\hbar c} > p_x = p_x \pm k \]
by multiplication of the two spinors. Hence we get
\[ \rho' = 2 e^2 V_0 \left[ \sum \frac{f_\xi (k) \bar{\psi}_\xi (k)}{k_\xi} + \sum \frac{f_- (k) \bar{\psi}_- (k)}{k_\xi} \right] \]
One can easily calculate* that the function
\[ f_\xi (k) \bar{\psi}_\xi (k) \]
vanishes for \( k = 0 \) and, if expanded in \( k \), has only terms of even power
\[ f_\xi (k) \bar{\psi}_\xi (k) = a_2 (p_x) k^2 + a_4 (p_x) k^4 + \cdots \]
Therefore because the second term in (12b) is identical to the first,
\[ \rho' = 4 e^2 V_0 \cos k \eta \sum \left[ a_2 (p_x) k^2 + a_4 (p_x) k^4 + \cdots \right] \]
Therefore,
\[ \rho' = \sqrt{A} \left[ A_2 k^2 + A_4 k^4 + \cdots \right] \]
where \( V \) is the original potential which we started with,
and
\[ A_\eta = 2 e^2 \sum a_\eta (p_x) = \frac{2 e^2}{(2\pi)^3} \int \rho^2 d\rho \ a_\eta (p_x) \]
There is no term in \( k^0 \) in the expansion (13) of \( \rho' \).
It would mean that a constant potential would give an effect, which is meaningless. This is also clear because if \( k = 0 \), \( p_\xi = p_x \)
and the matrix element is zero, so we are saved from this difficulty. Further, any terms with an odd value of \( k \) would be proportional to \( \sin k \eta \) giving a maximum effect where the potential \( V = V_0 \cos k \eta \) is zero: we do not expect this and symmetry rules it out. A symmetrical potential cannot give rise to a skew-symmetric charge density. We now calculate the expansion coefficients \( a_2, a_4, a_6 \) etc. If \( p_i \gg m c > \)
\[ f_\xi (k) \bar{\psi}_\xi (k) \approx \left( \frac{p_x / \hbar c - p_\xi / \hbar c}{|E_x| + |E_\xi|} \right)^2 c_2 \frac{k^2}{p_x} + c_4 \frac{k^4}{p_x} \]
* see later
Where $c_2$, $c_4$, etc., are numerical constants.

Hence in (14)

$$A_\hbar = \frac{c_\hbar}{\beta^* \hbar} \quad \text{for } \beta^* \gg m$$

Now we see, that, in the expression (13), the expansion coefficient $A_2$ of $P'$ will diverge logarithmically since

$$A_2 = \frac{2 e^2}{(2\pi)^3} \int \frac{p^3 dp}{p^2} \frac{c_\hbar}{p^3}$$

Remember now that our expression (15) for $a_n$ is only valid for $p \gg m c$; hence the lower limit of the integral (16) is not valid. The exact evaluation shows, that $a_n$ behaves for low $p$ like $\hbar$ constant. Hence we get $A_2 \sim \log (p_{\text{max}}/mc)$ where $p_{\text{max}}$ is the momentum at which we cut off the integration. All other $A_n$ are finite, since

$$A_\hbar = \frac{2 e^2}{(2\pi)^3} \int \frac{p^3 dp}{p^2} \frac{c_\hbar}{p^3}$$

which does not diverge at the upper limit for $\hbar > 2$.

What does the first term in the expansion (13) mean? It is proportional to

$$k^2 V = - \nabla^2 V \quad \text{for the cosine function assumed}$$

$$= \frac{4 \pi}{c} \times \rho \text{external}$$

This means that neglecting higher terms the induced charge density $\rho'$ is equal to a constant times the external charge, $\rho \text{external}$. Though unfortunately the constant of proportionality $A_2$ is infinite.

$$A_2 = \frac{e^2}{\hbar c} \log \left( \frac{p_{\text{max}}}{m c} \right)$$

$$= \frac{e^2}{\hbar c} \log \left( \frac{c}{a} \right)$$

where $a$ is the cut-off radius and $\chi_c$ is the Compton wavelength. $a_2$ is small for any reasonable value of the cut-off radius $a$. 5295/E
Further remarks on the effect of the outside field

We assumed an outside potential \( V = 2V_0 \cos kx \) and obtained for the induced charge due to polarization of the vacuum

\[
\rho' = A_2 k^2 V + A_4 k^4 V + \cdots \quad -(1)
\]

Now it is better to introduce the charge density \( \rho_o \) which produced the outside potential given by

\[
\nabla^2 V = -\rho_o \quad -(2)
\]

(We use rationalized units in which \( 4\pi \) becomes 1).

Therefore \( \rho_o = k^2 V_0 \) \( -(3) \)

Therefore \( \rho' = A_2 \rho_o + A_4 k^2 \rho_o + \cdots \quad -(4) \)

\[= A_2 \rho_o - A_4 \nabla^2 \rho_o + A_6 \nabla^4 \rho_o - \cdots \quad -(5)\]

i.e. the higher terms are proportional to derivatives of \( \rho_o \).

We can also write

\[
\rho'(x) = \int S'(x - x') \rho_o(x') \, dx' \quad -(6)
\]

This means, for example, that if \( \rho_o = \delta(x') \), \( \rho'(x) = S(x) \).
This means that a point charge induces a vacuum polarization charge spread according to the function \( S \).

The equivalence between the two formulae for \( \rho' \) can be seen as follows. Assume \( \rho_o \) is slowly varying, then by Taylor's expansion

\[
\rho' = \int S'(x - x') \rho_o(x') \, dx' = \rho_o(x) \int S'(x - x') \, dx' + \sum \frac{d^k \rho_o}{dx^k} \int x'^k S'(x - x') \, dx' \quad -(7)
\]
Only even derivatives are finite. Thus we see that the \( A_2 \), \( A_4 \), ..., of eq. (1) are the moments of \( S \). We have already seen that \( A_2 \) which is the zero-moment is infinite

\[
A_2 = \frac{e^2}{\kappa} \epsilon_0 \left( \kappa c / \alpha \right) \tag{8}
\]

while the others are finite

\[
A_4 = e^2 / \kappa c \times \text{some number} \tag{9}
\]

Hence \( \int S'/(-\epsilon')^d \epsilon' \) is infinite while the other moments are of order \( \kappa / mc \). Therefore the function \( S \) has the properties indicated in the figure below. \( S \) is, of course just the induced charge around the proton, and we know that it is negative. This tells us all we need to know about \( S \).

Figure 1

Width of wings \( \sim \kappa / \kappa c \)
Depth of wings \( \sim e^2 / \kappa c \)
Centre \( \sim e^2 / \kappa c \) \( \alpha^{-3} \) to give logarithmic divergence

A good sample function for \( S \) could be

\[
S' \sim \frac{e^2}{\kappa c} \frac{1}{\alpha^3} \alpha^{-\alpha / \kappa c}, \quad \kappa c = \alpha / mc \tag{10}
\]

This is not correct, but has all the right properties.

Although the integral \( \int_0^{\infty} S'/d\kappa \) diverges, the integral down to the proton radius, 1 fermi, gives only electronic charge \( 1/137 \), and therefore the effect is very small. This means that the total charge induced in the vacuum is infinite if I cut off at zero, but still very small in any reasonable case.
With no cut-off we get a positive charge inducing an infinite negative charge. This means that if \( a \) is too small the charge will kill itself.

**Charge Renormalization**

By a reasonable cut-off we mean \( a \ll \lambda_c \) but \( \log (\lambda_c/a) \ll 137 \). If, however \( \log (\lambda_c/a) \gtrsim 137 \), the charge is tending to kill itself and we call that an unreasonable cut-off.

With a reasonable cut-off what we really observe as the proton charge is the bare charge \( Z \) reduced by the vacuum polarization; this gives the observed charge \( Z' \). Our ideal point proton would therefore be surrounded by a charge distribution as indicated in Fig 2. The effect of the vacuum polarization is that the interaction between the electron and the proton is changed, because what we have been calculating here is the effect of the charged proton on the electron world. However it is customary to put the blame for the change in the interaction entirely on to the electronic charge. It is useful to say that the outside charge is still \( Z \) but that the electron charge is altered from \( e \) to \( e' \). If we do this we must also say that there is an increased attraction when the particles are close together because the electron is then seeing the bare proton which has a greater charge. Hence, for a reasonable cut-off we write

\[
e' = e (1 - 1/137 \cdot \log \frac{\lambda_c}{a})
\]

and remember also that the attraction is increased at short distances.
Further development due to Landau

Let us see where we get with an unreasonable cut-off. The first order calculation is now no good because it gives an induced charge greater than the outside charge. Now we had for the reasonable region

\[ \kappa = \left( \frac{e'}{e} \right)^2 = \left[ 1 - \frac{2e^2}{\lambda c} \log\left( \frac{\lambda c}{a} \right) \right] \]  \hspace{1cm} (11)

Landau regards this as the beginning of an expansion in powers of \( e^2 \). For the unreasonable region it is no longer valid to take the first terms of the expansion and we must make a guess at the complete function. For example the functions

\[ e^{2\kappa} \left\{ - \frac{2e^2}{\lambda c} \log\left( \frac{\lambda c}{a} \right) \right\} \]  \hspace{1cm} (12)

and

\[ \left\{ 1 + \frac{2e^2}{\lambda c} \log\left( \frac{\lambda c}{a} \right) \right\}^{-1} \]  \hspace{1cm} (13)

Both give equation (11) in first approximation. As the cut off distance \( a \) tends to zero both these functions give \( e' \rightarrow 0 \), instead of \( e' \rightarrow \infty \) as given by equation (11).

Therefore let us put

\[ \kappa = \left( \frac{e'}{e} \right)^2 = e^{2\kappa} \left\{ - \frac{2e^2}{\lambda c} \log\left( \frac{\lambda c}{a} \right) \right\} \]  \hspace{1cm} (14)

(the term \( 1/\kappa \) occurs in the exponential because we want the value of \( e^2/\kappa c \) and the experimental value of \( 1/137 \) is of course \( e^2/\kappa c \)).

This gives

\[ k \log k = - \frac{2}{137} \log\left( \frac{\lambda c}{a} \right) \]  \hspace{1cm} (15)
and the form of the function $k \log k$ is indicated in Fig. 3. To obtain the minimum value of $a$, (this is called $a^*$), we want $k \log k$ at its maximum negative value $1/e$.

Therefore

$$\log(\frac{a}{a^*}) = \frac{137}{2e} \quad -(16)$$

If $a < a^*$, $k$ has to be imaginary which is impossible. This means that electrodynamics with no cut-off is self-contradictory. Therefore there must be a cut-off larger than $a^*$, a result which we also obtain from self-energy considerations.

**Conclusion**

Electrodynamics is not a finished theory. There must be some as yet unrecognized phenomenon that explains the cut-off. Peaceful co-existence methods allow us to make calculations, but have not really solved the problem. We hope that a future theory which explains the cut-off will also show a connection between electrodynamics and nuclear matter, and explain why all particles have the same electric charge.