On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields.*

By S. Tomonaga

(Received May 17, 1946)

§ 1. The formalism of the ordinary quantum theory of wave fields.

Recently Yukawa has made a comprehensive consideration about the basis of the quantum theory of wave fields. In his article he has pointed out the fact that the existing formalism of the quantum field theory is not yet perfectly relativistic.

Let \( v(xyz) \) be the quantity specifying the field, and \( \lambda(xyz) \) denote its canonical conjugate. Then the quantum theory requires the commutation relations of the form:

\[
\{v(xyz), v(x'y'z')\} = \{\lambda(xyz), \lambda(x'y'z')\} = 0
\]

\[
\{v(xyz), \lambda(x'y'z')\} = i\hbar \delta(x-x') \delta(y-y') \delta(z-z'),
\]

but these have quite non-relativistic forms.

The equations (1) give namely the commutation relations between the quantities at different points \((xyz)\) and \((x'y'z')\) at the same instant of time \(t\). The concept "same instant of time at different points" has, however, a definite meaning only one specifies some definite Lorentz frame of reference. Thus this is not a relativistically invariant concept.

Further, the Schrödinger equation for the \(\phi\)-vector representing the state of the system has the form:

\[
\left(\vec{\mathcal{H}} + \frac{\hbar}{i} \frac{\partial}{\partial t}\right)\phi = 0,
\]


** \([A, B] = AB - BA\). We assume that the field obeys the Bose statistics. Our considerations apply also to the case of Fermi statistics.
where $\tilde{H}$ is the operator representing the total energy of the field which is given by the space integral of a function of $v$ and $\lambda$. As we adopt here the Schrödinger picture, $v$ and $\lambda$ are operators independent of time. The vector representing the state is in this picture a function of the time, and its dependence on $t$ is determined by (2).

Also the differential equation (2) is no less non-relativistic. In this equation the time variable $t$ plays a role quite distinguished from the space coordinates $x$, $y$ and $z$. This situation is closely connected with the fact that the notion of probability amplitude does not fit with the relativity theory.

As is well known, the vector $\psi$ has, as the probability amplitude, the following physical meaning: Suppose the representation which makes the field quantity $v(xyz)$ diagonal. Let $\psi[v'(xyz)]$ denote the representative of $\psi$ in this representation.* Then the representative $\psi[v'(xyz)]$ is called probability amplitude, and its absolute square

$$W[v'(xyz)] = \left| \psi[v'(xyz)] \right|^2$$

(3)
gives the relative probability of $v(xyz)$ having the specified functional form $v'(xyz)$ at the instant of time $t$. In other words: Suppose a plane** which is parallel to the $xyz$-plane and intercepts the time axis at $t$. Then the probability that the field has the specified functional form $v'(xyz)$ on this plane is given by (3).

As one sees, a plane parallel to the $xyz$-plane plays here a significant role. But such a plane is only defined by referring to a certain frame of reference. Thus the probability amplitude is not a relativistically invariant concept in the space-time world.

§ 2. Four-dimensional form of the commutation relations.

As stated above, the laws of the quantum theory of wave fields are

---

* We use the square blackets to indicate a functional. Thus $\phi[v'(xyz)]$ means that $\phi$ is a functional of the variable function $v'(xyz)$. When we use ordinary blackets ( ), as $\psi(v'(xyz))$, we consider $\psi$ as an ordinary function of the function $v'(xyz)$. For example: the energy density is written as $H(v'(xyz), \lambda'(xyz))$ and this is also a function of $x$, $y$ and $z$, whereas the total energy $H = \int H(v'(xyz), \lambda'(xyz))dv$ is a functional of $v(xyz)$ and $\lambda(xyz)$ and is written as $\tilde{H}[v'(xyz), \lambda'(xyz)]$.

** We call a three-dimensional manifold in the four-dimensional space-time world simply "surface".
usually expressed as mathematical relations between quantities having their meanings only in some specified Lorentz frame of reference. But since it is proved that the whole contents of the theory are of course relativistically invariant, it must be certainly possible to build up the theory on the basis of concepts having relativistic space-time meanings. Thus, in his consideration, Yukawa has required with Dirac to generalize the notion of probability amplitude so that it fits with the relativity theory. We shall now show below that the generalization of the theory on these lines is in fact possible to the relativistically necessary and sufficient extent. Our results are, however, not so general as expected by Dirac and by Yukawa, but are already sufficiently general in so far as it is required by the relativity theory.

Let us suppose for simplicity that there are only two fields interacting with each other. The case of more number of fields can also be treated in the same way. Let \( v_1 \) and \( v_2 \) denote the quantities specifying the fields. The canonically conjugate quantities be \( \lambda_1 \) and \( \lambda_2 \) respectively. Then between these quantities the commutation relations

\[
\begin{align*}
&[[v_r(xyst), v_s(x'y's't')] = 0 \\
&[[\lambda_r(xyst), \lambda_s(x'y's't')] = 0 \\
&[[v_r(xyst), \lambda_s(x'y's't')] = i\hbar \delta(x-x') \delta(y-y') \delta(z-z') \delta(\tau - \tau')
\end{align*}
\]

must hold. The \( \psi \)-vector satisfies the Schrödinger equation

\[
\left( \overline{H}_1 + \overline{H}_2 + \overline{H}_{12} + \frac{\hbar}{i} \frac{\partial}{\partial t} \right) \psi = 0.
\]

In this equation \( \overline{H}_1 \) and \( \overline{H}_2 \) mean respectively the energy of the first and the second field. \( \overline{H}_1 \) is given by the space integral of a function of \( v_1 \) and \( \lambda_1 \), \( \overline{H}_2 \) by the space integral of a function of \( v_2 \) and \( \lambda_2 \). Further, \( \overline{H}_{12} \) is the interaction energy of the fields and is given by the space integral of a function of both \( v_1, \lambda_1 \) and \( v_2, \lambda_2 \). We assume (i) that the integrand of \( \overline{H}_{12} \), i.e. the interaction-energy density, is a scalar quantity, and (ii) that the energy densities at two different points (but at the same instant of time) commute with each other. In general, these two facts follow from the single assumption: the interaction term in the Lagrangean does not contain the time derivatives of \( v_1 \) and \( v_2 \).

If this energy density is denoted by \( H_{12} \), then we have
\( H_{12} = \int H_{12} \, dx \, dy \, ds \) \hspace{1cm} (6)

As we adopt here the Schrödinger picture, the quantities \( \nu \) and \( \lambda \) in \( H_1 \)
\( H_2 \) and \( H_{12} \) are all operators independent of time.

Thus far we have merely summarized the well known facts. Now, as
the first stage of making the theory relativistic, we suppose the unitary
operator
\[
U = \exp \left\{ \frac{i}{\hbar} (H_1 + H_2) t \right\}
\]
and introduce the following unitary transformations of \( \nu \) and \( \lambda \), and the
corresponding transformation of \( \phi \):
\[
\begin{align*}
V_r &= U \nu U^{-1}, \quad A_r = U \lambda U^{-1} \\
\phi &= U \phi.
\end{align*}
\]
(8)

As stated above, \( \nu \) and \( \lambda \) in (5) are quantities independent of time. But
\( V \) and \( A \) obtained from them by means of (8) contain \( t \) through \( U \). Thus they depend on \( t \) by
\[
\begin{align*}
\hbar V_r &= V_r H_r - H_r V_r \\
\hbar A_r &= A_r H_r - H_r A_r
\end{align*}
\]
(9)

These equations must necessarily have covariant forms against Lorentz
transformations, because they are just the field equations for the fields when
they are left alone without interacting with each other.

Now, the solutions of these “vacuum equations”, the equations which
the fields must satisfy when they are left alone, together with the commu­
taion relations (4), give rise to the relations of the following forms:
\[
\begin{align*}
&[[V_r(xyst), V_s(x'y's't')] = A_r(x-x', y-y', s-s', t-t')] \\
&[[A_r(xyst), A_s(x'y's't')] = B_r(x-x', y-y', s-s', t-t')] \\
&[[V_r(xyst), A_s(x'y's't')] = C_r(x-x', y-y', s-s', t-t')]
\end{align*}
\]
(10)

where \( A_r, B_r \) and \( C_r \) are functions which are combinations of the so-called four-dimensional \( \delta \)-functions and their derivatives. One denotes usually these four dimensional \( \delta \)-functions by \( D_r(xyst) \), \( r=1, 2 \). They are defined by
On a Relativistically Invariant Formulation

\[ D_r(xyst) = \frac{1}{16\pi^3} \iiint \left\{ \frac{\varepsilon(k_x x + k_y y + k_z z + k_r t)}{ik_r} - \frac{\varepsilon(k_x x + k_y y + k_z z - k_r t)}{ik_r} \right\} dk_x dk_y dk_z \]  \hspace{1cm} (11)

with

\[ k_r = \sqrt{k_x^2 + k_y^2 + k_z^2 + k_r^2}, \]  \hspace{1cm} (12)

\( x_r \) being the constant characteristic to the field \( r \). It can be easily proved that these functions are relativistically invariant.*

Since (10) gives, in contrast with (4), the commutation relations between the fields at two different world points \((xyst)\) and \((x'y's't')\), it contains no more the notion of same instant of time. Therefore, (10) is sufficiently relativistic presupposing no special frame of reference. We call (10) fourdimensional form of the commutation relations.

One property of \( D(xyst) \) will be mentioned here: When the world point \((xyst)\) lies outside the light cone whose vertex is at the origin, then \( D(xyst) \) vanishes identically:

\[ D(xyst) = 0 \hspace{1cm} \text{for} \hspace{1cm} x^2 + y^2 + z^2 - c^2 t^2 > 0. \]  \hspace{1cm} (13)

It follows directly from (13) that, if the world point \((x'y's't')\) lies outside the light cone whose vertex is at the world point \((xyst)\), the right-hand sides of (10) always vanish. In words: Suppose two world points \( P \) and \( P' \). When these points lie outside each other's light cones, the field quantities at \( P \) and field quantities at \( P' \) commute with each other.

§ 3. Generalization of the Schrödinger equation.

Next we observe the vector \( \Phi \) obtained from \( \psi \) by means of the unitary transformation \( U \). We see from (5), (7) and (8) that this \( \Phi \), considered as

* Suppose that a surface in the \( k_x k_y k_z k_r \) k-space is defined by means of the equation \( k^2 = k_x^2 + k_y^2 + k_z^2 + k_r^2 \). Then this surface has the invariant meaning in this space, since \( k_x^2 + k_y^2 + k_z^2 - k_r^2 \) is invariant against Lorentz transformations. The area of the surface element of this surface is given by \( dS = \sqrt{\frac{\partial k}{\partial k_x}^2 + \left(\frac{\partial k}{\partial k_y}\right)^2 + \left(\frac{\partial k}{\partial k_z}\right)^2 - 1} \, dk_x \, dk_y \, dk_z \). Now, since \( dS \) has the invariant meaning, we can thus conclude that \( \frac{dk_x \, dk_y \, dk_z}{k} \) is an invariant, and this results that the function defined by (11) is invariant.
a function of $t$, satisfies

$$\left\{ \int H_{1t}(V_1(xyzt), A_1(xyzt), V_2(xyzt), A_2(xyzt)) \, dx \, dy \, ds + \frac{\hbar}{i} \frac{\partial}{\partial t} \right\} \Phi = 0. \quad (14)$$

One sees that $t$ plays also here a role distinguished from $x$, $y$ and $z$: also here a plane parallel to the $xyz$-plane has a special significance. So we must in some way remove this unsatisfactory feature of the theory.

This improvement can be attained in the way similar to that in which Dirac\(^{10}\) has built up the so-called many-time formalism of the quantum mechanics. We will now recall this theory.

The Schrödinger equation for the system containing $N$ charged particles interacting with the electromagnetic field is given by

$$\left\{ \overline{H}_{el} + \sum_{n=1}^{N} H_n(q_n, p_n, a(q_n)) + \frac{\hbar}{i} \frac{\partial}{\partial t} \right\} \psi = 0. \quad (15)$$

Here $\overline{H}_{el}$ means the energy of the electromagnetic field, $H_n$ the energy of the $n$-th particle. $H_n$ contains, besides the kinetic energy of the $n$-th particle, the interaction energy between this particle and the field through $a(q_n)$, $q_n$ being the coordinates of the particle and $a$ the potential of the field. $p_n$ in (15) means as usual the momentum of the $n$-th particle.

We consider now the unitary operator

$$u = \exp \left\{ \frac{i}{\hbar} \overline{H}_{el} t \right\} \quad (16)$$

and introduce the unitary transformation of $a$:

$$\mathcal{U} = u a u^{-1} \quad (17)$$

and the corresponding transformation of $\psi$:

$$\Phi = u \psi. \quad (18)$$

Then we see that $\Phi$ satisfies the equation

$$\left\{ \sum_{n} H_n(q_n, p_n, \mathcal{U}(q_n, t)) + \frac{\hbar}{i} \frac{\partial}{\partial t} \right\} \Phi = 0. \quad (19)$$

In contrast with $a$, which was independent of times (Schrödinger picture), $\mathcal{U}$
contains \( t \) through \( n \). To emphasize this, we have written \( t \) explicitly as argument of \( \mathcal{A} \). We can prove that \( \mathcal{A} \) satisfies the Maxwell equations in vacuum (accurately speaking, we need special considerations for the equation \( \text{div} \mathbf{E} = 0 \)).

The equation (19) is the starting point of the many-time theory. In this theory one introduces then the function \( \Phi(q_1 t_1, q_2 t_2, \ldots, q_n t_n) \) containing so many time variables \( t_1, t_2, \ldots, t_n \) as the number of the particles in place of the function \( \Phi(q, q_2, \ldots, q_n, t) \) containing only one time variable,∗ and suppose that this \( \Phi(q_1 t_1, q_2 t_2, \ldots, q_n t_n) \) satisfies simultaneously the following equations:

\[
\left\{ H_n(q_n, q_n', \mathcal{A}(q_n, t_n)) + \frac{\hbar}{i} \frac{\partial}{\partial t_n} \right\} \Phi(q_1 t_1, q_2 t_2, \ldots, q_n t_n) = 0 \tag{20}
\]

\( n = 1, 2, \ldots, N. \)

This \( \Phi(t_1, t_2, \ldots, t_n) \), which is a fundamental quantity in the many-time theory, is related to the ordinary probability amplitude \( \phi(t) \) by

\[
\phi(t) = \Phi(t_1, t_2, \ldots, t_n). \tag{21}
\]

Now, the simultaneous equations (20) can be solved when and only when the \( N^2 \) conditions

\[
(H_n H_n' - H_n' H_n) \Phi(q_1 t_1, q_2 t_2, \ldots, q_n t_n) = 0 \tag{22}
\]

are satisfied for all pairs of \( n \) and \( n' \). If the world point \((q_n t_n)\) lies outside the light cone whose vertex is at the point \((q_n' t_n')\), we can prove \( H_n H_n' - H_n' H_n = 0 \). As the result, the function satisfying (20) can exist in the region where

\[
(q_n - q_n')^2 - c^2 (t_n - t_n')^2 \geq 0 \tag{23}
\]

is satisfied simultaneously for all values of \( n \) and \( n' \).

According to Bloch(∗) we can give \( \Phi(q_1 t_1, q_2 t_2, \ldots, q_n t_n) \) a physical meaning when its arguments lie in the region given by (23). Namely

\[
\mathcal{W}(q_1 t_1, q_2 t_2, \ldots, q_n t_n) = |\Phi(q_1 t_1, q_2 t_2, \ldots, q_n t_n)|^2 \tag{24}
\]

gives the relative probability that one finds the value \( q_1 \) in the measurement of the position of the first particle at the instant of time \( t_1 \), the value \( q_2 \) in

∗ Here we suppose the representation which makes the coordinates \( q_1, q_2, \ldots, q_N \) diagonal. Thus the vector \( \Phi \) is represented by a function of these coordinates.
the measurement of the position of the second particle at the instant of
time $t_n$, ... and the value $g_n$ in the measurement of the position of the $N$-
th particle at the instant of time $t_N$.

This is the outline of the many-time formalism of the quantum
mechanics.

We will now return to our main subject. If we compare our equation
(14) with the equation (19) of the many-time theory, we notice a marked
similarity between these two equations. In (19) stands the suffix $n$, which
designates the particle, while in (14) stand the variables $x$, $y$ and $z$, which
designate the position in space. Further, $\Phi$ is a function of the $N$
independent variables $g_1, g_2, ..., g_N$ giving the position of the $n$-th particle, while $\Psi$ is a functional of the infinitely many "independent variables" $v_1(xyz)$ and $v_n(xyz)$ giving the fields at the position $(xyz)$. Corresponding to the sum $\sum H_n$ in (19) the integral $\int H dxdydz$ stands in (14). In this way, to the suffix $n$ in (19) which takes the values 1, 2, 3, ..., $N$
correspond the variables $x$, $y$ and $z$ which take continuously all values from
$-\infty$ to $+\infty$.

Such a similarity suggests us to introduce infinitely many time variables
$\epsilon_{sys}$, which we may call local time * each for one position $(xyz)$ in the space
as we have introduced $N$ time variables, particle times, $t_1, t_2, ..., t_N$
each for one particle. The only difference consist in that we use in our case
infinitely many time variables whereas we have used $N$ time variables in
the ordinary many-time theory.

Corresponding to the transition from the use of the function with one
time variable to the use of the function of $N$ time variables, we must now
consider the transition from the use of $\Psi(t)$ to the use of a functional $\Psi[\epsilon_{sys}]$
of infinitely many time variables $\epsilon_{sys}$.

We regard now $\epsilon_{sys}$ as a function of $(xyz)$ and consider its variation
$\delta_{sys}$ which differs from zero only in a small domain $V_0$ in the neighbourhood of the point $(x_0, y_0, z_0)$. We will define the partial differential coefficient of the functional $\Psi[\epsilon_{sys}]$ with respect to the variable $\epsilon_{sys}$ in the
following manner:

$$ \frac{\partial \Psi}{\partial \epsilon_{sys}} = \lim_{\epsilon_{sys} \to 0} \frac{\Psi[\epsilon_{sys} + \epsilon_{sys}] - \Psi[\epsilon_{sys}]}{\epsilon_{sys}} dx dy dz $$

$$ \tag{25} $$

* The notion of local time of this kind has been occasionary introduced by Stueckerberg. [6]
We then generalize (14), and regard
\[
\left( H(x, y, z, t) + \frac{\hbar}{i} \frac{\delta}{\delta t_{xyz}} \right) \Psi = 0
\]
(26)
the infinitely many simultaneous equations corresponding to the \(N\) equations (20), as the fundamental equations of our theory. In (26) we have written, for simplicity, \(H_1(x, y, z, t)\) in place of \(H_1(V(xyz, t), V_2(xyz, t), \ldots).\)

In general, when we have a function \(F(V, \lambda)\) of \(V\) and \(\lambda\), we will write simply \(F(x, y, z, t)\) for \(F(V(xyz, t_{xyz}), \lambda(xyz, t_{xyz}))\), or still simpler \(F(P)\) denoting the point with the coordinates \((xyz, t_{xyz})\). Thus \(F(P')\) means \(F(x', y', z', t')\) or, more precisely, \(F(V(x'y'z', t_{xyz}'), (x'y'z', t_{xyz}'))\).

We will now adopt the equation (26) as the basis of our theory. For \(V_1(P), V_2(P), \lambda_1(P)\) and \(\lambda_1(P)\) in \(H_1\), the commutation relations (10) hold, where \(D(xyz)\) has the property (13). As the consequence, we have
\[
H_{12}(P)H_{12}(P') - H_{12}(P')H_{12}(P) = 0
\]
(27)
when the point \(P\) lies a finite distance apart from \(P'\) and outside the light cone whose vertex is at \(P\). Further, from our assumption (ii) the relation (27) holds also when \(P\) and \(P'\) are two adjacent points approaching in a space-like direction. Thus our system of equations (26) is integrable when the surface defined by the equations \(t = t_{xyz}\) considering \(t_{xyz}\) as a function of \(x, y\) and \(z\), is space-like.

In this way, a functional of the variable surface in the space-time world is determined by the functional partial differential equations (26). Corresponding to the relation (21) in case of many-time theory, \(\Psi[t_{xyz}]\) reduces to the ordinary \(\Psi(\xi)\) when the surface reduces to a plane parallel to the \(xyz\)-plane.

The dependent variable surface \(t = t_{xyz}\) can be of any (space-like) form in the space-time world, and we need not presuppose any Lorenz frame of reference to define such a surface. Therefore, this \(\Psi[t_{xyz}]\) is a relativistically invariant concept. The restriction that the surface must be space-like makes no harm since the property that a surface is space-like or time-like does not depend on a special choice of the reference system. It is not necessary, from the standpoint of the relativity theory, to admit also time-like surfaces for the variable surface, what was required by Dirac and by Yukawa. Thus we consider that \(\Psi[t_{xyz}]\) introduced above is already the sufficient generalization of the ordinary \(\phi\)-vector, and assume that the quantum-
theoretical state* of the fields is represented by this functional vector.

Let $C$ denote the surface defined by the equation $t = t_{\text{rel}}$. Then $\mathcal{F}$ is a functional of the surface $C$. We write this as $\mathcal{F}[C]$. On $C$ we take a point $P$, whose coordinates are $(xyz, t_{\text{rel}})$, and suppose a surface $C'$ which overlap $C$ except in a small domain about $P$. We denote the volume of the small world lying between $C$ and $C'$ with $d\omega_p$. Then we may write (25) also in the form:

$$\frac{\delta \mathcal{F}[C]}{\delta C_p} = \lim_{{\epsilon \to 0}} \frac{\mathcal{F}[C'] - \mathcal{F}[C]}{d\omega_p}. \quad (28)$$

Then (28) can be written in the form:

$$\left\{ H_{12}(P) + \frac{i}{\hbar} \frac{\delta}{\delta C_p} \right\} \mathcal{F}[C] = 0. \quad (29)$$

This equation (29) has now a perfect space-time form. In the first place, $H_{12}$ is a scalar according to our assumption (i); in the second place, the commutation relations between $V(P)$ and $A(P)$ contained in $H_{12}$ has the four-dimensional forms as (10), and finally the differentiation $\frac{\delta}{\delta C_p}$ is defined by (28) quite independently of any frame of reference.

A direct conclusion obtained from (29) is that $\mathcal{F}[C']$ is obtained from $\mathcal{F}[C]$ by the following infinitesimal transformation:

$$\mathcal{F}[C'] = \left\{ 1 - \frac{i}{\hbar} H_{12}(P) d\omega_p \right\} \mathcal{F}[C]. \quad (30)$$

When there exist in the space-time world two surfaces $C_1$ and $C_2$ a finite distance apart, we need only to repeat the infinitesimal transformations in order to obtain $\mathcal{F}[C_2]$ from $\mathcal{F}[C_1]$. Thus

$$\mathcal{F}[C_2] = \frac{\hbar}{\alpha} \left\{ 1 - \frac{i}{\hbar} H_{12}(P) d\omega_p \right\} \mathcal{F}[C_1] \quad (31)$$

The meaning of this equation is as follows: We devide the world region lying between $C_1$ and $C_2$ in small elements $d\omega_p$ (it is necessary that each world element is surrounded by two space-like surfaces). We consider for

each world element the infinitesimal transformation $1 - \frac{i}{\hbar} H_{\text{rel}}(P) \text{d}\omega$. Then
we take the product of these transformations, the order of the factor being
taken from $C_1$ to $C_2$. This product transforms then $\mathcal{F}[C_1]$ into $\mathcal{F}[C_2]$.

The surfaces $C_1$ and $C_2$ must be here both space-like, but otherwise
they may have any form and any configuration. Thus $C_1$ does not necessarily
lie afterward against $C_2$; $C_1$ and $C_2$ may even cross with each other.

The relation of the form (31) has been already introduced by Heisenberg. It can be
regarded as the integral form of our generalized Schrödinger equation (29).


We must now find the physical meaning of the functional $\mathcal{F}[C]$. As
regards this we can make a similar consideration as Bloch has done for the
case of ordinary many-time theory. Besides the fact that in our case there
appear infinitely many time variables, one point differs from Bloch's case
that in (16) the unitary operator $u$ is commutable with the coordinates $q_1,
q_2, \ldots, q_m$ our $U$ is not commutable with the field quantities $v_1(xyz)$ and $v_2(xyz)$. Noting this difference and treating the continuum infinity as
the limit of an ennumerable infinity by some artifice, for instance, by the procedure
of Heisenberg and Pauli, Bloch's consideration can be applied also here
almost without any alteration. We shall give here only the results.

Let us suppose that the fields are in the state represented by a vector
$\mathcal{F}[C]$. We suppose that we make measurements of a function $f(v_1, v_2, \lambda_1, \lambda_2)$
at every point on a surface $C_1$ in the space-time world. Let $P_1$ denote the
variable point on $C_1$, then, if $f(P_1)$ at any two "values" of $P_1$ commute
with each other, the measurement of $f$ at each of these two points do not
interfere with each other. Our first conclusion says that in this case the
expectation value of $f(P_1)$ is given by

$$\langle f(P_1) \rangle = \langle \mathcal{F}[C_1], f(P_1) \mathcal{F}[C_1] \rangle$$

(32)

where $f(P_1)$ means $f(V_1(P_1), \ldots)$ according to our convention on page
35, and the symbol $\langle (A, B) \rangle$ with double blackets is the scalar product of
two vectors $A$ and $B$. It is impossible in case of continuously many
degree of freedom to represent this scalar product by an integral of the
product of two functions. For this purpose we must replace the continuum
infinity by an at least ennumerable infinity.
More generally, we suppose a functional $F[f(P')]$ of the independent variable function $f(P')$, regarding $f(P')$ as a function of $P'$. Then the expectation value of this $F$ is given by

$$F[f(P')] = \langle \Psi[C_1], F[f(P')] \rangle[C_1] \rangle. \quad (33)$$

A physically interesting $F$ is the projective operator $M[v_1'(P_1), v_2'(P_1), V_1(P_1), V_2(P_1)]$ belonging to the "eigen-value" $v_1'(P_1), v_2'(P_1)$ of $V_1(P_1), V_2(P_1)$. Then its expectation value

$$M[v_1'(P_1), v_2'(P_1); V_1(P_1), V_2(P_1)]$$

$$= \langle \Psi[C_1], M[v_1'(P_1), v_2'(P_1); V_1(P_1), V_2(P_1)] \rangle[C_1] \rangle \) (34)$$

gives the probability that the field 1 and the field 2 have respectively the functional form $v_1'(P_1)$ and $v_2'(P_1)$ on the surface $C_1$. As $C_1$ is assumed to be space-like, the measurement of the functional $M$ is possible (the measurements of $V_1(P_1)$ and $V_2(P_1)$ at all points on $C_1$ mean just the measurement of $M$).

Thus far we have made no mention of the representation of $\Psi[C]$. We use now the special representation in which $V_1(P_1)$ at all points on $C_1$ are simultaneously diagonal. It is always possible to make all $V_1(P_1)$ and $V_2(P_1)$ diagonal when the surface $C_1$ is space-like. In this representation $\Psi[C_1]$ is represented by a functional $\Psi[v_1'(P_1), v_2'(P_1); C_1]$ of the eigenvalues $v_1'(P_1)$ and $v_2'(P_1)$ of $V_1(P_1)$ and $V_2(P_1)$. The projection operator $M$ has in this representation such diagonal form that (34) is simplified as follows

$$W[v_1'(P_1), v_2'(P_1)] = \langle \Psi[v_1'(P_1), v_2'(P_1); C_1] \rangle[C_1] \rangle.$$

In this sense we can call $\Psi[v_1'(P_1), v_2'(P_1); C_1]$ "generalized probability amplitude".

§ 5. Generalized transformation functional.

We have stated above that between $\Psi[C_1]$ and $\Psi[C_2]$ the relation (31) holds, where $C_1$ and $C_2$ are two space-like surfaces in the space-time world. We see thus that the transformation operator
On a Relativistically Invariant Formulation

\begin{equation}
T[C_2; C_1] = \frac{\theta_2}{\alpha} \left(1 - \frac{i}{\hbar} H \right) \alpha \omega_0
\end{equation}

plays an important role. It is evident that also this operator has a space-
time meaning.

Similarly as the special representative of the \( \phi \)-vector, the probability
amplitude, has a distinct physical meaning, there is a special representation
in which the representative of the transformation operator \( T[C_2; C_1] \) has
a distinct physical meaning.

We introduce namely the mixed representative of \( T[C_2; C_1] \) whose
rows refer to the representation in which \( V_1(P) \) and \( V_2(P) \) at all points
on \( C_1 \) become diagonal and whose column refer to the representation in
which \( V_1(P) \) and \( V_2(P) \) at all points on \( C_2 \) become diagonal. We denote
this representation by

\begin{equation}
\{ v_1''(P_2), v_2''(P_2) \mid T[C_2; C_1] \mid v_1'(P_1), v_2'(P_1) \}, \quad (37)^* 
\end{equation}

or simpler:

\begin{equation}
\{ v_1''(P_2), v_2''(P_2) \mid v_1'(P_1), v_2'(P_1) \}. \quad (38)^* 
\end{equation}

If we note here the relation (35), we see that we can give the matrix
elements of this representation the following meaning: One measures the
field quantities \( V_1 \) and \( V_2 \) at all points on \( C_2 \) when the fields are prepared
in such a way that they have certainly the values \( v_1'(P) \) and \( v_2'(P) \) at
all points on \( C_1 \). Then

\begin{equation}
W[v_1''(P_2), v_2''(P_2) ; v_1'(P_1), v_2'(P_1)] \\
= [v_1''(P_2), v_2''(P_2) \mid v_1'(P_1), v_2'(P_1)]^2 \quad (39)
\end{equation}

gives the probability that one obtains the result \( v_1''(P_2) \) and \( v_2''(P_2) \) in this
measurement. In this proposition we have assumed that \( C_2 \) lies afterward
against \( C_1 \).

From this physical interpretation we may regard the matrix element
(37), or (38), considered as a functional of \( v_1''(P_2), v_2''(P_2) \) and \( v_1'(P_1), v_2'(P_1) \), as
the generalization of the ordinary transformation function \( \langle q_2'' \mid q_1' \rangle \).

* As the matrix elements are functionals of \( \tau(P) \) we use here the square brackets.
As a special case it may happen that \( C_s \) lies apart from \( C_t \) only in a portion \( S_s \) and a portion \( S_t \) of \( C_t \) and \( C_t \) respectively, the other parts of \( C_t \) and \( C_t \) overlapping with each other (see Fig. 1).

In this case the matrix elements of \( T[C_s; C_t] \) depend only on the values of the fields on the portions \( S_s \) and \( S_t \) of the surfaces \( C_t \) and \( C_t \). In this case we need for calculating \( T[C_s; C_t] \) to take the product in (36) only in the closed domain surrounded by \( S_s \) and \( S_t \), thus

\[
T[S_s; S_t] = \frac{\hbar}{i} \left( 1 - e^{-\frac{i}{\hbar} H dt} \right).
\]

The matrix elements of the mixed representation of this \( T \) is a functional of \( v_{1}'(\rho_1), v_{2}'(\rho_2) \) and \( v_{1}''(\rho_2), v_{2}''(\rho_2) \) where \( \rho_1 \) denotes the moving point on the portion \( S_s \), and \( \rho_2 \) the moving point on the portion \( S_t \). This matrix is independent on the field quantities on the other portions of the surfaces \( C_t \) and \( C_t \).

The matrix element of \( T[S_s; S_t] \) regarded as a functional of \( v_{1}'(\rho_1), v_{2}'(\rho_1) \) and \( v_{1}''(\rho_1), v_{2}''(\rho_1) \) has the properties of g. t. f. (generalized transformation functional) of Dirac. But in defining our g. t. f. we had to restrict the surfaces \( S_s \) and \( S_t \) to be space-like, while Dirac has required his g. t. f. to be defined also referring to the time-like surfaces. As mentioned above, however, such a generalization as required by Dirac is superfluous so far as the relativity theory concerns.

It is to be noted that for the physical interpretation of \( [v_{1}''(P_t), v_{2}''(P_t)] \)

\[
[v_{1}'(P_t), v_{2}'(P_t)]
\]

it is not necessary to assume \( C_t \) to lie afterward against \( C_t \). Also when the inverse is the case, we can as well give the physical meaning for \( W \) of (39): One measures the field quantities \( V_1 \) and \( V_t \) at all points on \( C_t \) when the fields are prepared in such a way that they would have certainly the values \( v_{1}'(P_t) \) and \( v_{2}'(P_t) \) at all points on \( C_t \) if the fields were left alone until \( C_t \) without being measured before on \( C_t \). Then \( W \) gives the probability that one finds the results \( v_{1}''(P_t) \) and \( v_{2}''(P_t) \) in this measurement on \( C_t \).
§ 6. Concluding remark.

We have thus shown that the quantum theory of wave fields can be really brought into a form which reveals directly the invariance of the theory against Lorentz transformations. The reason why the ordinary formalism of the quantum field theory is so unsatisfactory lies in the fact that one has built up this theory in the way which is too much analogous to the ordinary non-relativistic mechanics. In this ordinary formalism of the quantum theory of fields the theory is divided into two distinct sections: the section giving the kinematical relations between various quantities at the same instant of time, and the section determining the causal relations between quantities at different instants of time. Thus the commutation relations (1) belong to the first section and the Schrödinger equation (2) to the second.

As stated before, this way of separating the theory into two sections is very unrelativistic, since here the concept “same instant of time” plays a distinct role.

Also in our formalism the theory is divided into two sections. But now the separation is introduced in another place: In our formalism the theory consists of two sections, one of which gives the laws of behavior of the fields when they are left alone, and the other of which gives the laws determining the deviation from this behavior due to interactions. This way of separating the theory can be carried out relativistically.

Although in this way the theory can be brought into more satisfactory form, no new contents are added thereby. So, the well known divergence difficulties of the theory are inherited also by our theory. Indeed, our fundamental equations (29) admit only catastrophic solutions as can be seen directly in the fact that the unavoidable infinity due to non-vanishing zero-point amplitudes of the fields inheres in the operator $H_{II}(P)$. Thus, a more profound modification of the theory is required in order to remove this fundamental difficulty.

It is expected that such a modification of the theory would possibly be introduced by some revision of the concept of interaction, because we meet no such difficulty when we deal with the non-interacting fields. This revision would then result that in the separability of the theory into two sections, one for free fields and one for interactions, some uncertainty would be introduced. This seems to be implied by the very fact that, when we formulate the quantum field theory in a relativistically satisfactory manner,
this way of separation has revealed itself as the fundamental element of the theory.

Physics Department,
Tokyo Bunrika University.

References
(1) H. Yukawa: Kagaku, 12 (1924), 251, 282 and 322.
(2) P. A. M. Dirac: Phys. Z. USSR., 3 (1933), 64.
(3) W. Pauli: Solvey Berichte (1939).