THE CONNECTION BETWEEN OPERATORS IN THE HEISENBERG AND INTERACTION REPRESENTATIONS IN LOCAL QUANTUM FIELD THEORY

A. D. SUKHANOV

L. Ya. Karpov Physico-chemical Institute

Submitted to JETP editor January 17, 1964

J. Exptl. Theoret. Phys. (U.S.S.R.) 47, 1303-1314 (October, 1964)

The problem of establishing the correct connection between the renormalized Heisenberg-representation operators and the free operators is considered in the framework of perturbation theory. For this purpose we introduce the concept of two types of T-products (Wick and Dyson products), in which the contractions differ by quasi-local terms in the case of derivatives of fields. Then in any local field theory, although the S matrix can be expressed in terms of either of the T-products with simultaneous replacement of $L_I^{in}(x)$ by $-H_I^{in}(x, \sigma)$, there are two possibilities for the definition of the Heisenberg operators. The possibility chosen here corresponds to a nonunitary connection between the operators in question and assures that the renormalized Heisenberg operator is independent of σ. We also propose a correct method for determining the matrix $S(\sigma, -\infty)$ and the interaction Hamiltonian in the interaction representation.

1. STATEMENT OF THE PROBLEM

In recent years there has been a rapid development of axiomatic methods for constructing quantum field theory (see in particular [1,2]). A number of well known successes have been achieved in this way, but at present it still cannot be asserted that we know about all of the peculiarities of the mathematical apparatus of the theory. Further study of this apparatus in the framework of any one system of axioms is in practice an extremely complicated task. Owing to this there is a tendency[3] toward using in such investigations concepts from various systems of axioms, and also toward bringing in, within reasonable limits, apparatus borrowed from the Lagrangian formalism. Along with this, in this sort of research there is, as before, insufficient use of the apparatus of the Hamiltonian formalism; this is in all probability due to a widespread prejudice against the fundamental quantity of this formalism—the "half-way" S matrix. In a series of studies which we have made we have attacked the problem of showing that this prejudice is to a large extent without any foundation. At the same time the use of the Hamiltonian formalism in the study of the mathematical structure of field theory allows the elucidation of a number of complex questions which do not yield readily to study by other methods.[1)]

As is well known, the basis of the covariant Hamiltonian formalism[4] is an assumption taken over from nonrelativistic quantum mechanics, that there exists an interaction representation, the operators and state vectors of which are connected with the analogous quantities in the Heisenberg representation by a unitary transformation which makes the average values equal:

$$\mathbf{F}(x) = S^+(\sigma, -\infty) F^{\text{int}}(x; \sigma) S(\sigma, -\infty), \tag{1}$$

$$\Psi^{\text{int}}(\sigma) = S(\sigma, -\infty)\Psi(-\infty), \tag{2}$$

with

$$i\delta S(\sigma, -\infty) / \delta\sigma(x) = H_I^{\text{int}}(x; \sigma) S(\sigma, -\infty). \tag{3}$$

Furthermore one of the essential postulates of the theory is that the Heisenberg-representation operator $\mathbf{F}(x)$, as an observable quantity which is therefore truly local, is independent of the spacelike surface σ.

The Dyson matrix[5] or "half-way" S matrix $S(\sigma, -\infty)$ which realizes the transformations (1) and (2) must satisfy the following general require-

1)The first communication about these results was contained in the author's report at the All-union Conference on Colliding-beam Accelerators and the Physics of High-energy Particles (Novosibirsk, June 1963).

ments: 1) relativistic covariance; 2) independence of the particular choice of σ; 3) unitarity:

$$S(\sigma, -\infty)S^{+}(\sigma, -\infty) = S^{+}(\sigma, -\infty)S(\sigma, -\infty) = 1; \quad (4)$$

4) the group property:

$$S(\infty, \sigma)S(\sigma, -\infty) = S(\infty, -\infty) = S, \quad (5)$$

where the operator S is the scattering matrix; 5) the boundary conditions

$$\lim_{\sigma\to\infty} S(\sigma, -\infty) = S, \qquad \lim_{\sigma\to-\infty} S(\sigma, -\infty) = 1; \quad (6)$$

6) finiteness.

Up to now it has been possible in perturbation theory to obtain an expression for the matrix $S(\sigma, -\infty)$ which satisfies only the requirements 1)–5). At the same time, the question of the finiteness of the matrix $S(\sigma, -\infty)$ has remained open even in renormalizable theories, because it turned out[6] that, unlike the S matrix, the formally obtained expression for $S(\sigma, -\infty)$ contains besides the "ultraviolet" divergences also special "surface" divergences (even for finite regularization masses M_i^2). Our earlier investigations[7] have shown, however, that at least for fixed M_i^2 the problem of "surface" divergences in the matrix $S(\sigma, -\infty)$ can be successfully solved. Owing to this the obtaining of a finite matrix $S(\sigma, -\infty)$ in perturbation theory now does not seem such a hopeless task.

A more serious argument against the existence of the matrix $S(\sigma, -\infty)$ has arisen in the framework of the axiomatic method. Namely, on the basis of the Källén-Lehmann theorem[8] we have the following formula for the renormalized operator of a scalar field in the Heisenberg representation

$$\langle 0|[\dot{\mathbf{A}}(x), \mathbf{A}(y)]|_{x^0=y^0}|0\rangle = -iZ_3^{-1}\delta(\mathbf{x}-\mathbf{y}),$$

$$Z_3^{-1} = \int_0^\infty \rho(m^2)\,dm^2 > 1. \quad (7)$$

Meanwhile, if we choose as the operator corresponding to $\mathbf{A}(x)$ in the interaction representation the operator $\varphi^{\mathrm{in}}(x)$ of a free scalar field,[9] which satisfies the conditions

$$(\Box_x - m^2)\,\varphi^{in}(x) = 0,$$

$$[\dot{\varphi}^{in}(x), \varphi^{in}(y)]|_{x^0=y^0} = -i\delta(\mathbf{x}-\mathbf{y}), \quad (8)$$

where m^2 is the physical mass, we get an obvious contradiction between the formulas (1), (8), and (7), because there is no unitary transformation that can turn $Z_3^{-1} > 1$ into unity.

This contradiction is usually[10] interpreted to mean that the connection between the renormalized operator $\mathbf{A}(x)$, which satisfies (7), and the free operator $\varphi^{\mathrm{in}}(x)$, which satisfies (8), must actually be of a nonunitary nature, and therefore, it is said, a finite and unitary "half-way" S matrix cannot exist in quantum field theory.

Meanwhile Yang and Feldman,[9] in discussing the example of a theory of the neutral vector field, had already actually suggested, but not used, a different way to liquidate this contradiction. In fact, although the connection between the operators $\mathbf{A}(x)$ and $\varphi^{\mathrm{in}}(x)$ may be of a nonunitary nature, we cannot conclude from this at all that the matrix $S(\sigma, -\infty)$ does exist, but only that the operator in the interaction representation that corresponds to $\mathbf{A}(x)$ according to (1) is not identical with the free operator $\varphi^{\mathrm{in}}(x)$ that satisfies (8).

This idea was developed to some extent in papers by Takahashi and Umezawa,[11] and was recently used again by Kirzhnits[12] in nonlocal field theory. What he suggested was that we separate the problem of obtaining a finite matrix $S(\sigma, -\infty)$ and that of the connection between the field operators $\mathbf{A}(x)$ and $\varphi^{\mathrm{in}}(x)$, and think of this connection in the explicitly nonunitary form

$$\mathbf{A}(x) = S^{+}(\sigma, -\infty)\varphi^{in}(x)S(\sigma, -\infty) + \chi(x; \sigma), \quad (9)$$

where $\chi(x; \sigma)$ is an operator which depends on the point x and is a functional of σ, and the matrix $S(\sigma, -\infty)$ is obtained not in the usual way[4,5] from (3), but directly from the S matrix.

In the present paper we shall develop a formal mathematical apparatus which enables us within the framework of perturbation theory on one hand to define the matrix $S(\sigma, -\infty)$ and the Hamiltonian $H_I^{\mathrm{in}}(x; \sigma)$ more concretely than is usual, and on the other hand to introduce a nonunitary connection between an arbitrary renormalized Heisenberg operator and the corresponding free operator in local quantum field theory. In subsequent papers these results will be applied to concrete Heisenberg operators.[2)]

2)Before going on to expound the main results, we make only one further remark. All of our further arguments are made on the assumption that in the expressions used a Pauli-Villars regularization has been carried out and that all of the M_i^2 are fixed. It will be seen from what follows that actually there remain no grounds for supposing that in the limit $M_i^2 \to \infty$ the matrix $S(\sigma, -\infty)$ cannot be made finite in at least some sense, but we shall for the time being defer the practical realization of the passage to the limit, and also the proof of the existence of the matrix $S(\sigma, -\infty)$ within the framework of the purely axiomatic method.

2. THE TWO TYPES OF CHRONOLOGICAL PRODUCTS

Already in [13] we called attention to a fact which has been known in principle but not sufficiently appreciated. Namely, in theories with derivative couplings (or vector fields)—and because of the presence of counter terms in the effective interaction Lagrangian $L_I^{in}(x;1)$ [14] this includes practically all local field theories—for a particular S matrix one formally uses always the same expression (the T-exponential) in two functional arguments which are different in such a case—the Lagrangian $L_I^{in}(x)$ and the Hamiltonian $H_I^{in}(x;\sigma)$. In this connection it was shown that the apparent contradiction between the two expressions for the S matrix can be eliminated if we understand that actually we are acquainted with not one, but two different (of course, for identical arguments) types of chronological products, which we shall hereafter call the Wick product (T_W) and the Dyson product (T_D).

The difference between them is that the contractions in the T_W-product are actually defined in the momentum representation, for example for a scalar field [14]

$$D^c(x) = \frac{1}{(2\pi)^4}\int \frac{e^{ikx}dk}{m^2-k^2-i\varepsilon}, \tag{10}$$

whereas the contractions in the T_D-product are defined in the coordinate representation, and this gives

$$D^c(x) = \theta(x^0)D^-(x) - \theta(-x^0)D^+(x). \tag{11}$$

For scalar fields themselves these two expressions are identical (on some class of regular functions), but for chronological products of derivatives of scalar fields there is a difference of quasilocal terms with definite coefficients, because in the T_W-product, in accordance with the requirements of Wick's theorem, the derivatives act on $D^c(x)$ as a whole and are simply applied to the exponential, so that we have equations of the type

$$\left\langle 0 \left| T_W\left(\frac{\partial\varphi^{in}(x)}{\partial x^\alpha}\frac{\partial\varphi^{in}(y)}{\partial y^\beta}\right)\right| 0\right\rangle = \frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial y^\beta}\langle 0 \mid T_W(\varphi^{in}(x)\varphi^{in}(y)) \mid 0\rangle, \tag{12}$$

and this is so also for $x = y$, whereas in the T_D-product the derivatives act only on the D^- and D^+ functions and not on the θ functions.[3)]

The difference between the T_W- and T_D-products is a special case of the well known possibility [14] of defining the chronological product for equal arguments while keeping the S matrix unchanged.[4)] In particular, the following special cases are of interest in principle for what follows:

$$\left\langle 0 \left| T_W\left(\frac{\partial\varphi^{in}(x)}{\partial x^\alpha}\frac{\partial\varphi^{in}(y)}{\partial y^\beta}\right)\right| 0\right\rangle = \left\langle 0 \left| T_D\left(\frac{\partial\varphi^{in}}{\partial x^\alpha}\frac{\partial\varphi^{in}}{\partial y^\beta}\right)\right| 0\right\rangle + in_\alpha n_\beta\delta(x-y), \tag{13}$$

$$\langle 0 \mid T_W(\Box_x\varphi^{in}(x)\Box_y\varphi^{in}(y)) \mid 0\rangle = \langle 0 \mid T_D(\Box_x\varphi^{in}(x)\Box_y\varphi^{in}(y)) \mid 0\rangle + i(\Box_x+m^2)\delta(x-y), \tag{14}$$

where n_α is the unit vector normal to the surface σ through the point $x = y$. It can be seen from (13) that by its definition the T_W-product is independent of σ, whereas the T_D-product in the general case will depend on σ. This is due to the fact that the expression (10) actually does not depend on any surface σ, whereas (11) involves θ functions, which because of relativistic covariance must be written $\theta(\sigma)$, where σ is an arbitrary spacelike surface. The result is that differentiation with respect to particular components gives rise to an explicit dependence on σ in the T_D-product. At the same time it can be seen from (14) that the T_W-product can contain quasi-local terms with derivatives in coordinate space.

We call attention to the fact that formulas of the type of (13) and (14) are algebraic relations which are externally extremely reminiscent of the formula for passage from ordinary products to normal products according to Wick's theorem. Indeed, for this case we can establish a remarkable analog of Wick's theorem, according to which every T_W-

3)Particular attempts to deal with the different expressions for the contractions in calculating the S matrix in theories with derivatives have been made in [15,11], but the point of view was not applied consistently and was not developed further. Recently a treatment of this problem much like ours has been given in [16] for the case of contractions of vector fields, but the exposition in that paper contains a number of inaccuracies and mistakes.[17] Outside the framework of perturbation theory the T_W-product has actually been used in [18], and recently (together with the T_D-product) in [3].

4)A deeper reason for the difference is evidently that when contractions with derivatives are formed there is a contribution to the corresponding spectral representations from the integral over a large circle, which is included in one type of T-product and not in the other.[3]

product can be expressed in terms of a set of T_D-products with all possible "quasi-contractions." These latter will be quasi-local terms of the type of those shown in the right members of (13) and (14), and of course will differ in accordance with the number of derivatives and the types of fields. We shall give a general expression for these terms later.

In concluding our examination of the differences between T_W- and T_D-products, we shall observe how these differences show up in the S matrix when it is put in the form of a functional expansion in normal products of the fields $\varphi^{in}(x)$, as is usually done [1,2] in constructing the matrix by the axiomatic method. In this connection we call attention to the fact that the derivatives of the field operators can be written in the following form

$$\frac{\partial\varphi^{in}(x)}{\partial x^0} = \int\limits_{-\infty}^{\infty} \frac{\partial}{\partial x^0}\,\delta(x-y)\,\varphi^{in}(y)\,dy, \qquad (15)$$

and because of this the question arises as to which of the arguments, x or y, should be used for the chronological ordering. If it is made with the argument y, this corresponds to the T_W-product, and use of x corresponds to the T_D-product. It is clear that when we have to do with an expansion of the S matrix in the operators $\varphi^{in}(x)$ alone (without derivatives)—that is, when its coefficient functions include also terms with derivatives of δ functions—then chronological ordering between the S matrix and any other operators, such as we shall be making here, will lead to a T_W-product, in which one first orders chronologically and then differentiates. In the opposite case we arrive at a T_D-product.

3. THE T-PRODUCT IN THE S MATRIX AND THE DEFINITION OF THE MATRIX $S(\sigma, -\infty)$

Historically, the first way [4,5] to obtain the matrix $S(\sigma, -\infty)$ was to solve the Tomonaga-Schwinger equation (3). In [7] we showed that in a local renormalizable theory one can in principle obtain by this method an expression for the "half-way" S matrix which satisfies all of the requirements stated in Sec. 1 (for $M_I^2 = \text{const}$), but associated with this there is the necessity of obtaining the Hamiltonian $H_I^{in}(x; \sigma)$ by an independent method, in theories with higher derivatives—and because of the presence of counterterms this means practically all known theories. At the same time, the usual ways [19] of obtaining $H_I^{in}(x, \sigma)$ lead to the correct expression only in special cases, when the Lagrangian involves only first powers of first derivatives of the fields. In particular, these ways do not enable us to bring out possible added terms in the Hamiltonian which do not depend on σ. Only the method for obtaining $H_I^{in}(x; \sigma)$ proposed in [9] and developed in [20] allows us to obtain the expression for it in a broader class of theories, but this method also, while extremely complicated, is still not exhaustive.

As for the "half-way" S matrix, it has been shown in [7] that the expression

$$S(\sigma, -\infty) = \lim_{g\to\theta_\sigma} S(g), \qquad (16)$$

where $S(g)$ is the effective scattering matrix, [14] will, after a "surface" renormalization has been carried out, satisfy all of the conditions imposed on $S(\sigma, -\infty)$ (for $M_I^2 = \text{const}$). And once the existence of such an expression has been proved, it can also be obtained directly from the complete S matrix by following definite rules. This approach is the most convenient to take for the definition of the "half-way" S matrix, as it avoids the complicated procedure of "surface" renormalization.

We shall start from the expression for the S matrix proposed in [14], which is naturally to be written in terms of the T_W-product in the form

$$S = T_W \exp\left\{ i \int\limits_{-\infty}^{\infty} L_I{}^{in}(x; 1)\,dx \right\}, \qquad (17)$$

where $L_I^{in}(x; 1)$ is the effective interaction Lagrangian [14] including counterterms

$$L_I{}^{in}(x; 1) = L_I{}^{in}(x) + \sum_{n=1} \int\limits_{-\infty}^{\infty} \Lambda_{n+1}^{in}(x, y_1, \ldots, y_n)\,dy_1\ldots dy_n. \qquad (18)$$

The simplest way to get the matrix $S(\sigma, -\infty)$ from an S matrix of the form (17) is to "cut" it into two "halves" by taking some surface σ as the limit of integration for each variable—that is, by constructing the quantity

$$T_W \exp\left\{ i \int\limits_{-\infty}^{\sigma} L_I{}^{in}(x; 1)\,dx \right\}. \qquad (19)$$

As explained in [7], a necessary preliminary step is to remove all of the integrations over y_i in (18); this does not destroy the finiteness of the S matrix, but prevents the appearance in $S(\sigma, -\infty)$ of "surface" divergences associated with the presence of counterterms Λ_{n+1}^{in} which contain derivatives.

In this case, however, even in renormalizable theories the expression (19), as has already been pointed out in [21], will not only contain "surface" divergences (for $M_I^2 = \text{const}$) but will also not satisfy the group property (5) and the unitarity condition (4). This last is not hard to understand if we recall that the definition of contractions in the T_W-product includes quasi-local terms with derivatives [cf. (14)], which for the same reason as applies to the counterterms Λ^{in}_{n+1} will lead in the expression (19) to nonintegrable products of the type $\delta(\sigma - x^0) \times \delta(\sigma - x^0)$ and to loss of the properties (5) and (4). This is actually the way the fact that the T_W-product (unlike the T_D-product) is nonunitary manifests itself; we shall discuss this in more detail below.

To avoid this difficulty, the natural first proposal is to transform the S matrix in the form (17) identically from a T_W-product to a T_D-product, using for this purpose the analog of Wick's theorem which we formulated above. This transformation does not disturb any of the properties of the S matrix, including its finiteness. Moreover, a number of properties of the complete S matrix become more understandable. The result obtained by the transformation is

$$S = T_W \exp\left\{ i \int_{-\infty}^{\infty} L_I{}^{in}(x;\, 1)\, dx \right\}$$

$$\equiv T_D \exp\left\{ -i \int_{-\infty}^{\infty} H_I{}^{in}(x;\, \sigma)\, dx \right\}. \qquad (20)$$

The formally introduced expression for the operator $H_I^{in}(x; \sigma)$ is the most natural one to adopt as the definition of the interaction Hamiltonian in the interaction representation. In other words, we give the name of the Hamiltonian $H_I^{in}(x; \sigma)$ to the "Lagrangian" with which the S matrix is expressed in terms of a T_D-product. Then, if the S matrix satisfies the usual requirements,[14] in particular the condition of unitarity, the $H_I^{in}(x; \sigma)$ so defined will satisfy the conditions of: a) relativistic covariance, b) hermiticity, c) integrability, and d) finiteness (for $M_i^2 = \mathrm{const}$).

Having made the transformation (20), we now define the "half-way" S matrix by the relation

$$S(\sigma, -\infty) = T_D \exp\left\{ -i \int_{-\infty}^{\sigma} H_I{}^{in}(x';\, \sigma')\, dx' \right\}, \qquad (21)$$

from which we can obtain $H_I^{in}(x; \sigma)$ directly by the formula inverse to (3). In the next section we consider the question: in what sort of local field theories will the matrix $S(\sigma, -\infty)$ in the form (21) satisfy (for $M_i^2 = \mathrm{const}$) all of the requirements that we have imposed above?—and also the question of the most general expression for $H_I^{in}(x; \sigma)$.

It must also be pointed out that whereas previously[13] we emphasized only the fact that "surface" terms (involving normals) and terms with "surface" divergences are in principle different in their operator structures (in renormalizable theories) and must not be confused, it is now clear that both types of terms are, speaking generally, different aspects of the same phenomenon, associated with the derivation of correct expressions for the matrix $S(\sigma, -\infty)$ and for $H_I^{in}(x; \sigma)$. In particular, if the Lagrangian contains only first derivatives, then already $H_I^{in}(x; \sigma)$ is not equal to $-L_I^{in}(x)$ and contains terms which depend through the n_α on the concrete choice of the surface σ, but not only the expression (21) but also (19) will still satisfy all of the requirements imposed on $S(\sigma, -\infty)$. If, on the other hand, the Lagrangian contains derivatives of higher orders, then not only is $H_I^{in}(x; \sigma) \neq -L_I^{in}(x)$, but also there is a dependence on the mere fact that the surface σ is introduced, and owing to this the group property and unitarity are maintained only for $S(\sigma, -\infty)$ in the form (21). At the same time, the property of being independent of the choice of the surface σ is safely maintained only for the expression (19), as a result of the properties of the T_W-product which we have noted, and this must be taken into account in what follows.

4. THE INTERACTION HAMILTONIAN IN LOCAL QUANTUM FIELD THEORY

We have dealt with the problem of obtaining the Hamiltonian $H_I^{in}(x; \sigma)$ by the method just indicated, in the case when there are first and second derivatives in the first degree in the Lagrangian, in [13] and [21], and obtained results which agree with the results of other methods for getting $H_I^{in}(x; \sigma)$.[9,19,20] Here we shall consider this problem in the general case.

For this it is first necessary to know the expression for the "quasi-contractions" in the case of an arbitrary number of derivatives. These expressions can be obtained if we consider the Fourier transforms of the T_D-products in question and use the formula for covariant separation of a four-vector into transverse and longitudinal parts,

$$k_\alpha = k_\alpha{}^{\mathrm{long}} + k_\alpha{}^{\mathrm{trans}}, \quad k_\alpha{}^{\mathrm{trans}} = n_\alpha (nk). \qquad (22)$$

Then, if for convenience we introduce the "quasi-contractions" in the momentum representation by the formula

$$\left\langle 0 \left| T_W \left(\frac{\partial^n \varphi^{in}(x)}{\partial x_{\alpha_1} \ldots \partial x_{\alpha_n}} \varphi^{in}(y) \right) \right| 0 \right\rangle$$

$$- \left\langle 0 \left| T_D \left(\frac{\partial^n \varphi^{in}(x)}{\partial x_{\alpha_1} \ldots \partial x_{\alpha_n}} \varphi^{in}(y) \right) \right| 0 \right\rangle$$

$$= \Delta_{\alpha_1\ldots\alpha_n}(x-y) = \frac{1}{(2\pi)^4}\int e^{ik(x-y)}\Delta_{\alpha_1\ldots\alpha_n}(k)\,dk, \quad (23)$$

then in a straightforward way we get for the $\Delta\alpha_1 \ldots \alpha_n(k)$ (beginning with the second derivative)

$$\Delta_{\alpha\beta}(k) = i^3 n_\alpha n_\beta, \quad (24a)$$

$$\Delta_{\alpha\beta\gamma}(k) = i^4[k_\alpha n_\beta n_\gamma + k_\beta n_\alpha n_\gamma + k_\gamma n_\alpha n_\beta - 2n_\alpha n_\beta n_\gamma (nk)], \quad (24b)$$

$$\Delta_{\alpha\beta\gamma\delta}(k) = i^5\{[k_\alpha k_\beta n_\gamma n_\delta + k_\alpha k_\gamma n_\beta n_\delta + k_\alpha k_\delta n_\beta n_\gamma + k_\beta k_\gamma n_\alpha n_\delta + k_\beta k_\delta n_\alpha n_\gamma + k_\gamma k_\delta n_\alpha n_\beta] - 2(nk)[k_\alpha n_\beta n_\gamma n_\delta + k_\beta n_\alpha n_\gamma n_\delta + k_\gamma n_\alpha n_\beta n_\delta + k_\delta n_\alpha n_\beta n_\gamma] + n_\alpha n_\beta n_\gamma n_\delta [m^2 - k^2 + 4(nk)^2]\}. \quad (24c)$$

It is then easy to apply mathematical induction, but we shall not do this owing to lack of space.

As will be seen from what follows, the most important special cases of formulas of the type (24) are expressions in which covariant summation is taken over either all pairs of α_i or over all pairs of indices except one or two. These formulas can be obtained by making the summations in (24a)—(24c) and going further by induction. They are as follows:

$$\Delta_s(k) = i^{n+1}\theta_{n2}\sum_{j=0}^{s-1}(m^2)^j(k^2)^{s-1-j}, \quad (25a)$$

$$\Delta_{\alpha.\,s}(k) = i^{n+1}\theta_{n2}k_\alpha\sum_{j=0}^{s-1}(m^2)^j(k^2)^{s-1-j}, \quad (25b)$$

$$\Delta_{\alpha\beta,\,s-1}(k) = i^{n+1}\theta_{n2}\left\{(m^2)^{s-1}n_\alpha n_\beta + \sum_{j=0}^{s-2}(m^2)^j(k^2)^{s-2-j}k_\alpha k_\beta\right\}, \quad (25c)$$

where $s = [n/2]$ and the symbol Δ_s means that s d'Alembertians have been formed as the result of covariant summation of $[n/2]$ pairs of α_i.

But knowing the expressions for the "quasi-contractions" still does not solve the problem of deriving the general expression for $H_I^{in}(x;\sigma)$. In practice there are a number of difficulties, which it is simplest to explain with examples. As one of these we consider the theory with $L_I^{in}(x) = g : [\varphi^{in}(x)]^4 :$. Then

$$L_I{}^{in}(x,1) = gZ_4 : [\varphi^{in}(x)]^4 : + \frac{1}{2}[Z_3 - 1]\{ : \varphi^{in}(x)\Box_x\varphi^{in}(x) : - m^2 : [\varphi^{in}(x)]^2 : \}, \quad (26)$$

where in perturbation theory Z_3 and Z_4 are sums of terms, each of which diverges logarithmically for $M_I^2 \to \infty$. This example is sufficiently general for the renormalizable theories, since (26) occurs in the Lagrangian both of scalar electrodynamics and of pseudoscalar meson theory.

If we apply in this case the analog of Wick's theorem, it turns out[21] that there is a contribution to $H_I^{in}(x;\sigma)$ not only from the treatment of $T_W[L_I^{in}(x;1)\,L_I^{in}(y;1)]$, as in the case of the bare $L_I^{in}(x)$ for scalar electrodynamics,[13] but also from multiple products of Lagrangians, $T_W[L_I^{in}(x;1)\,L_I^{in}(y;1)\,L_I^{in}(z;1)]$ and so on. Thus to obtain the expression for $H_I^{in}(x;\sigma)$ it is necessary to carry out a summation of a large number of quasi-local terms by applying rather complex combinatory procedures. Furthermore it can be shown that in this case the additional terms in $H_I^{in}(x;\sigma)$ will have the same operator structure as the terms in $L_I^{in}(x;1)$. This means that they do not depend on σ and cannot be brought out, as in [19], by starting from the integrability condition (see also [20]).

In particular it is interesting to obtain the part of $H_I^{in}(x;\sigma)$ that has the operator structure $:\varphi^{in}(x)\Box_x\varphi^{in}(x):$. It is of the form

$$-\tilde{H}_I{}^{in}(x;\sigma) = \frac{1}{2}(Z_3-1):\varphi^{in}(x)\Box_x\varphi^{in}(x): - \left(\frac{Z_3-1}{2}\right)^2\frac{3}{2!}:\varphi^{in}(x)\Box_x\varphi^{in}(x): + \left(\frac{Z_3-1}{2}\right)^3\frac{15}{3!}:\varphi^{in}(x)\Box_x\varphi^{in}(x): + \ldots = (1 - Z_3^{-1/2}):\varphi^{in}(x)\Box_x\varphi^{in}(x):. \quad (27)$$

The summation in (27) is of course rather formal, but, firstly, we are taking $M_I^2 = \text{const}$, and then Z_3 is finite, and, secondly, we inevitably have to resort to such summations when we deal with renormalization constants in perturbation theory.[5,14] As for the coefficients in the series, the denominators are obvious, and the numerators are determined by a direct counting up of the number of quasi-local terms.

The next important peculiarity of the derivation of $H_I^{in}(x;\sigma)$ is that in theories with higher derivatives the direct change (20) from a T_W-product to a T_D-product gives rise to products of "quasi-contractions" with identical arguments, which give contributions to $H_I^{in}(x;\sigma)$ of the type $i\delta^4(0)$, which are infinite (for $M_I^2 = \text{const}$), and in addition are nonhermitian. In renormalizable theories such expressions appear only in the vacuum terms, and therefore can be ignored altogether. In nonrenormalizable theories, however, since the unitarity of the S matrix in the form (20) is a direct consequence of the hermiticity of $H_I^{in}(x;\sigma)$, such contributions to $H_I^{in}(x;\sigma)$ will lead to violation of unitarity. If now we wish to preserve the unitarity of the S matrix, then for this it is necessary to introduce corresponding nonhermitian and infinite (for $M_i^2 = \text{const}$) additions to the Lagrangian so that they will cancel out when the transformation (20) is made. An analogous situation has already been observed in nonlocal theory.[22] It follows that if we start from a unitary S matrix, then in deriving $H_I^{in}(x; \sigma)$ we must include in it only quasi-local terms obtained from $T_W[L_I^{in}(x_1) \ldots L_I^{in}(x_n)]$ with single "quasi-contractions" of an adjacent pair of Lagrangians.

Finally, we must call attention to the fact that if $L_I^{in}(x)$ involves higher derivatives taken with respect to separate components, then the expression for $H_I^{in}(x; \sigma)$ will be a purely formal one, since in this case the group property (and also unitarity) cannot be made compatible with the condition that $S(\sigma, -\infty)$ be independent of the concrete choice of σ. In fact, as long as we have to do with the "quasi-contraction" (24a) we can always return in $S(\sigma, -\infty)$ from the T_W-product to the T_D-product, but this is impossible in the cases (24b) and (24c). Therefore it is reasonable to search for a $H_I^{in}(x; \sigma)$ only in theories in which $L_I^{in}(x)$ contains the first derivatives with respect to separate components, but the d'Alembertian can still occur to any power. Owing to this the "quasi-contractions" (25a)—(25c) are most interesting, and $H_I^{in}(x; \sigma)$ cannot involve more than two normals n_α.[13] In particular, vector electrodynamics is not a theory of this type.[5)]

Going over from scalar fields to fields of other types, we note that since $S^c(x-y)$ contains only the first derivative, for spinor fields themselves $H_I^{in}(x; \sigma) = -L_I^{in}(x)$. From the point of view of the derivation of $H_I^{in}(x; \sigma)$ the vector field is equivalent to the first derivative of a scalar field, and the derivatives of vector (and spinor) fields correspond to higher derivatives of a scalar field.

5. DETERMINATION OF THE RENORMALIZED OPERATORS IN THE HEISENBERG REPRESENTATION

As has already been noted, the possibility of a nonunitary connection between corresponding operators in local field theory follows from a number of results in both nonlocal[12] and also local[9-11,12] theory. Moreover, in this case a nonlocal connection is not only possible, but also necessary, because the assumption that the connection between the operators $\mathbf{F}(x)$ and $F^{in}(x)$ is unitary is in contradiction with the requirement that the renormalized operator $\mathbf{F}(x)$ in the Heisenberg representation be independent of σ.

As the definition of the connection between the renormalized operator in the Heisenberg representation and the free operator in the interaction representation we choose the following expression:

$$\mathbf{F}(x) = S^+ T_W(F^{in}(x) S). \tag{28}$$

The guiding considerations in favor of this formula are given by the manner of definition of Heisenberg operators, which is independent of the Hamiltonian formalism and goes back to Schwinger.[24] In this procedure one starts from the Lagrangian formalism and takes as the foundation an S matrix of the form (17), after which one adds to the Lagrangian (and not to the Hamiltonian!) appropriate classical functions and takes the variation with respect to them. Then there is naturally obtained a formula of the form (28) with the T_W-product.

The advantage of such a definition is primarily that in any theory the operator $\mathbf{F}(x)$ in the form (28) is surely independent of σ, because, as was pointed out above, precisely the T_W-product is independent of σ, whereas the T_D-product in general does depend on σ.

At the same time it is obvious that it is precisely for the T_D-product that there is the identity[6)]

$$T_D(F^{in}(x) S) \equiv S(\infty, \sigma) F^{in}(x) S(\sigma, -\infty), \tag{29}$$

where the S matrix is of the form (20) and $S(\sigma, -\infty)$ is of the form (21), which automatically leads to an expression of the type (1)

$$\mathbf{F}'(x, \sigma) = S^+(\sigma, -\infty) F^{in}(x) S(\sigma, -\infty) \tag{30}$$

instead of (28).

Thus there is the following interesting situation. As long as we are dealing with a theory without derivatives, then first, the definitions of T_W- and T_D-products coincide, and second, $H_I^{in}(x; \sigma) = -L_I^{in}(x)$. The result is that the expressions (17) and (20) for the S matrix coincide, and so also do the expressions (28) and (30) for Heisenberg operators, so that the connection between $\mathbf{F}(x)$ and $F^{in}(x)$ is both unitary and preserves lack of dependence on the surface.

If, on the other hand, the theory involves derivatives, then although the S matrix itself is not affected by this, the formulas which before were

5)It is interesting to note that in the attempt to get a $H_I^{in}(x; \sigma)$ in such a theory,[16] starting from a Hermitian $L_I^{in}(x)$, only the fact that the additional terms that arise are infinite was noted, and not the fact that they are not Hermitian.

6)It is understood throughout that the surface σ passes through the point x.

single are split up, and there are two cases:

1. T_W-products: the formula (17) for the S matrix; formula (28) for $\mathbf{F}(x)$, preserving lack of dependence on σ.

2. T_D-products: formula (20) for the S matrix; formula (30) for $\mathbf{F}'(x, \sigma)$, preserving the unitary character of the connection.

The definition (28) can also be extended to operators which could not be obtained in the original formalism of Schwinger.[24] In particular, we can determine in this way the free Lagrangian in the Heisenberg representation, and from the sum $\Lambda_{tot}(x) = \Lambda_0(x) + \Lambda_I(x)$ we can obtain the equations of motion. Then, however, there arises the complicated problem of expressing an arbitrary Heisenberg operator $\mathbf{F}(x)$ in terms of the renormalized field operators $\mathbf{A}(x)$, since the T_W-product, unlike the T_D-product, does not have the property of being multiplicative.

To emphasize the seriousness of the difference between the T_W- and T_D-products, we point out one further fact. The counter terms of the proper energy which go into the interaction Lagrangian $L_I^{in}(x; 1)$ and the current $j^{in}(x)$ are not equal to zero (because of the free field equations) only when under the sign of the nonunitary T_W-product. Under the sign of the T_D-product, on the other hand, they are identically equal to zero, so that their inclusion does no good. Of course all of this also applies to the expressions for the S matrix and the matrix $S(\sigma, -\infty)$.

Owing to the nonunitarity of the T_W-product the expression (28) for the renormalized Heisenberg operator is not too convenient for further investigation. Besides this, the assumption that the matrix $S(\sigma, -\infty)$ exists obliges us to assume that in the interaction representation there exists another operator $F^{int}(x; \sigma)$, which in general is not equal to $F^{in}(x)$ [see (1)], and that the connection between this other operator and $\mathbf{F}(x)$ is unitary. Our next aim will be to reduce the expression (28) to a form convenient for use by finding operators $F^{int}(x; \sigma)$ corresponding to the various operators $\mathbf{F}(x)$. The basic apparatus in this research will be the analog of Wick's theorem established above, which allows us to express the T_W-product in terms of the T_D-product, and a subsidiary tool will be the expression (21) for the matrix $S(\sigma, -\infty)$ in terms of $H_I^{in}(x; \sigma)$.

In conclusion I express my sincere gratitude to B. V. Medvedev, on whose initiative this work was undertaken. I also express my gratitude to D. V. Shirkov, V. Ya. Faĭnberg, D. A. Kirzhnits, and A. V. Astakhov for fruitful criticism.

[1] Lehmann, Symanzik, and Zimmermann, Nuovo cimento **1**, 205 (1955); **6**, 319 (1957).

[2] Bogolyubov, Medvedev, and Polivanov, Voprosy teorii dispersionnykh sootnosheniĭ (Problems in the Theory of Dispersion Relations), Fizmatgiz, 1958.

[3] B. V. Medvedev and M. K. Polivanov, Lectures at the International School of Theoretical Physics, Dubna, March, 1964.

[4] S. Tomonaga, Prog. Theoret. Phys. **1**, 27 (1946). J. Schwinger, Phys. Rev. **74**, 1439 (1948).

[5] F. J. Dyson, Phys. Rev. **75**, 486 (1949).

[6] E. C. G. Stueckelberg, Phys. Rev. **81**, 130 (1951).

[7] A. D. Sukhanov, Candidate's Dissertation, Joint Institute for Nuclear Research, 1962; DAN SSSR **145**, 1042 (1962), Soviet Phys. Doklady **7**, 734 (1963); JETP **43**, 932 (1962), Soviet Phys. JETP **16**, 660 (1963).

[8] G. Källén, Helv. Phys. Acta **25**, 417 (1952). H. Lehmann, Nuovo cimento **11**, 342 (1954).

[9] C. N. Yang and D. Feldman, Phys. Rev. **79**, 972 (1950).

[10] R. Haag, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. **29**, No. 12, (1955).

[11] Y. Takahashi and H. Umezawa, Prog. Theoret. Phys. **9**, 14, 501 (1953). H. Umezawa, Quantum Field Theory, North-Holland–Interscience, 1956.

[12] D. A. Kirzhnits, JETP **41**, 551 (1961), Soviet Phys. JETP **14**, 395 (1962); JETP **45**, 143 (1963), Soviet Phys. JETP **18**, 103 (1964).

[13] A. D. Sukhanov, JETP **41**, 1915 (1961), Soviet Phys. JETP **14**, 1361 (1962).

[14] N. N. Bogolyubov and D. V. Shirkov, Vvedenie v teoriyu kvantovannykh poleĭ (Introduction to Quantum Field Theory), Gostekhizdat, 1957.

[15] Z. Koba, Prog. Theoret. Phys. **5**, 139 (1950).

[16] T. D. Lee and C. N. Yang, Phys. Rev. **128**, 885 (1962).

[17] A. D. Sukhanov, JETP **44**, 2087 (1963), Soviet Phys. JETP **17**, 1464 (1963).

[18] D. A. Slavnov and A. D. Sukhanov, JETP **41**, 1940 (1961), Soviet Phys. JETP **14**, 1379 (1962). B. V. Medvedev, DAN SSSR **153**, 313 (1963), Soviet Phys. Doklady **8**, 1092 (1964).

[19] P. T. Matthews, Phys. Rev. **76**, 1657 (1949). S. Kaneszwa and Z. Koba, Prog. Theoret. Phys. **4**, 297 (1949).

[20] Y. Katayama, Prog. Theoret. Phys. **10**, 31 (1953).

[21] A. D. Sukhanov, JETP **43**, 1400 (1962), Soviet Phys. JETP **16**, 992 (1963).

[22] B. V. Medvedev, JETP **32**, 87 (1957), Soviet Phys. JETP **5**, 48 (1957).

[23] V. Ya. Faĭnberg, JETP **40**, 1758 (1961), Soviet Phys. JETP **13**, 1237 (1961). A. D. Sukhanov, DAN SSSR **145**, 545 (1962), Soviet Phys. Doklady **7**, 603 (1963).

[24] J. Schwinger, Proc. Natl. Acad. Sci. U.S. **37**, 452 (1951). Utiyama, Sunakawa, and Imamura, Prog. Theoret. Phys. **8**, 77 (1952).

Translated by W. H. Furry
190