# On The Statistical Interpretation of Quantum Theory

R. L. STRATONOVICH
*Moscow State University*

(Submitted to JETP editor October 12, 1956)
J. Exptl. Theoret. Phys. (U.S.S.R.) 32, 1483-1495 (June, 1957)

We consider various principles of operator arrangements as basis for a statistical interpretation. An extremely concise expression for the distribution of trajectories in functional space is obtained, using the action functional, for the case of the ground state. The bichronological arrangement in time is a Markov process, and a continuous distribution density is found for it.

THERE ARE THREE STAGES in the probability interpretation of quantum theory. The first stage is the usual interpretation of the absolute square of the wave function as a probability density. The second stage is the search for a joint distribution for physical quantities, corresponding to non-commuting operators, at a given time.[1,2] If we want to consider the evolution in time, it is not sufficient to limit ourselves to a calculation of the change with time of these distributions. Moreover, it is of some general interest to study joint multiple distributions which correspond to different times and which take into account the correlation between states at different times. The consideration of such distributions forms the third stage of the problem mentioned at the beginning, and we shall consider it here from first principles. Thus the subject of the investigation of the present paper will be closely related to a paper by Feynman,[3] but in contradistinction to the latter our work is based upon the idea that the shape of the distribution should follow simply from an accepted "principle of arranging" operators. This fact plays a systematic role, and makes it in particular possible to show a relation between the papers of Feynman[3] and Moyal[1] and also to indicate another possible space-time interpretation.

## 1. PRINCIPLES OF OPERATOR ARRANGEMENTS AND THEIR ROLE IN A STATISTICAL INTERPRETATION

One can easily determine the probability density distribution $w(a)$ of a physical quantity $a$ corresponding to an operator $A$, if one uses the quantum rule for averaging. For a given state described by a wave function $\Psi$ (or, more generally, by any matrix density) we have

$$w(a) = \langle \delta(A-a)\rangle \equiv (\Psi, \delta(A-a)\Psi). \quad (1)$$

This definition is connected with the considerations that the usual statistical method of averaging $\int F(a)\, w(a)\, da$ should agree with the quantum average $\langle F(A)\rangle$.

The evaluation of expression (1) can, of course, be reduced to calculating the averages of the exponential function or of powers,

$$w(a) = \frac{1}{2\pi}\int e^{-iaI}\langle e^{iAI}\rangle\, dI = \frac{1}{2\pi}\int e^{-iaI}\sum\frac{(iI)^n}{n!}\langle A^n\rangle\, dI. \quad (2)$$

The integral of the exponent can be interpreted as the integral showing how to treat the characteristic function $\langle(\exp(iAI)\rangle$.

One can prove that expression (1) is non-negative, but this is not important for our discussion. The fact is that for a comprehensive statistical interpretation of quantum theory (that is, for a consideration of joint distributions of quantities corresponding to non-commuting operators) one must apparently inevitably allow the possibility of local negative values of the distribution density, without this, however, leading to negative probabilities for integral operations which are connected to physical measurements. This should still be a small price to pay for the possibility to use for typical quantum phenomena the classical presentation and the usual probability theory. Unfortunately, the situation is to a large degree still complicated because of the involved nature of the distributions; we shall return to this below.

The generalization of Eq. (1) for the case of non-commuting operators $A_1, A_2, \ldots$ has the form

$$w(a_1, a_2, \ldots) = \langle \delta(A_1 - a_1)\,\delta(A_2 - a_2)\ldots\rangle, \quad (3)$$

but now the question arises in what order the operators under the average sign will act. The determination of this order in Eq. (3) is equivalent to the

determination of the order of operation of the operators in the characteristic function $\langle \exp i(A_1 I_1 + A_2 I_2 + \dots)\rangle$ or in the compound moments $\langle A_{n_1} A_{n_2} \dots\rangle$.

We denote by Pr any principle of operator arrangement. To be precise, let Pr fix the coefficients $c(i, \dots, l)$ in the expansion

$$\Pr A_{n_1} \dots A_{n_s} = \sum c(i, \dots, l)\, A_{n_i} \dots A_{n_l} \quad (s = 2, 3, \dots) \tag{4}$$

for all possible permutations of the operators $A$ ($i, \dots, l$ is a permutation of the numbers $1, \dots, s$). These coefficients must satisfy the conditions

$$\sum c(i, \dots, l) = 1; \quad c(i, \dots, l) \geqslant 0. \tag{4a}$$

The operators under the Pr-sign can be rearranged. The action on a function of operators is defined by the action on each term of its Taylor expansion.

If we choose any one ordering principle Pr, we can use it to determine the characteristic function of a multiple distribution

$$\Theta(I_1, I_2, \dots) = \langle \Pr \exp i(A_1 I_1 + A_2 I_2 + \dots)\rangle, \tag{5}$$

the distribution density $\langle \Pr \delta(A_1 - a_1)\, \delta(A_2 - a_2) \dots\rangle$, correlation functions *etc.*, *i.e.*, on the basis of this principle we can construct a complete statistical picture.

The difficulty connected with the large number of possible choices for Pr is mitigated by the fact that in actual cases one can show that there are only a small number of specific rules for arranging operators which should especially be preferred (because of symmetry, relativistic invariance, or other reasons). One cannot, however, attain a sufficiently simple principle, and one must take into account several "preferable" statistical pictures and their interconnection. We shall consider below separately the various cases of distribution principles and the distribution laws connected with them.

The simplest ordering principle occurs if in all products $A_{n_1} A_{n_2} \dots$ one of the operators acts always either before or after another, and on the right hand side of Eq. (4) there is only one term. One can then assign to the operators Feynman's index $s$ (which is such that a smaller value of $s$ corresponds to earlier action).[4] In particular, if we identify $s$ with the time $t$, we shall have the chronological arrangement Pr = P. The chronological arrangement leads to a simple result in the Heisenberg representation, when all the operators taken at one particular time commute (the non-commutability of the coordinates and the momenta is not taken into account as one can express the momenta as derivatives with respect to the coordinates).

It is convenient to denote the reverse way of arranging the operators ($s = -t$) by Pr = P*. An advantage of the chronological and antichronological arrangements is their simplicity and relativistic invariance; the disadvantage is that real functions of Hermitian operators will, generally speaking, correspond to non-Hermitian operators. Also the probability density corresponding to them is complex.[5]

One can avoid this defect by going over to the arrangement principle

$$\Pr^{(1)} = \tfrac{1}{2}(\mathrm{P} + \mathrm{P}^*). \tag{6}$$

Another corrected mode of expansion, which is also based upon the P and P* arrangements, can most conveniently be formulated by its application to the exponential function

$$\Pr^{(2)} \exp \sum A_i u_i = \mathrm{P}^* \left[\exp \tfrac{1}{2} \sum A_i u_i\right] \mathrm{P} \left[\exp \tfrac{1}{2} \sum A_i u_i\right]. \tag{7}$$

A natural generalization of this arrangement principle is given by the equation

$$\Pr^{(2n)} \exp \sum A_i u_i = \left\{\mathrm{P}^* \left[\exp \tfrac{1}{2n} \sum A_i u_i\right] \mathrm{P} \left[\exp \tfrac{1}{2n} \sum A_i u_i\right]\right\}^n. \tag{8}$$

In the limit $n \to \infty$ we get the completely symmetrical ordering

$$\Pr^{(s)} \exp \sum A_i u_i = \exp \sum A_i u_i. \tag{9}$$

From Eqs. (7) to (9) one can determine the operation of Pr on other functions, for instance by differentiating with respect to $u_{n_1}, u_{n_2}, \dots$, and subsequently putting these arguments equal to zero we can find $\Pr A_{n_1} A_{n_2} \dots$.

Below we shall consider the distributions and characteristic functionals, which we shall mark by an index in brackets, corresponding to the above forms of arranging operators.

The symmetrical arrangement (9) is convenient for considering linear transformations of operators. It was used in Ref. 1 for a statistical interpretation of quantum states at a fixed time. By applying it to

the coordinate and momentum operators one obtained in that paper the Wigner distribution, which could, however, equally easily have been obtained from the $\mathrm{Pr}^{(2)}$ arrangement.

Among the ordering principles (7) to (9) introduced by us the bichronological principle $\mathrm{Pr}^{(2)}$, which we shall denote by R, plays a special role as it is bilinear in P, just as the quantum average value is bilinear in the wave function. For a state described by the vector $\psi$, the characteristic functional $\langle \mathrm{R} \exp \int A(t)u(t)\,dt \rangle$ is the absolute square of the vector $\mathrm{P}[\exp \int Au\,dt/2]\psi$. As a counterpart to R we have the principle R′ corresponding to a characteristic functional $\Theta^{(\mathrm{R}')}$ which is the square of the vector $\mathrm{P}^*[\exp \int Au\,dt/2]\psi$. In the majority of cases $\Theta^{(\mathrm{R}')}$ and $\Theta^{(\mathrm{R})}$ are connected according to the equation $\Theta^{(\mathrm{R}')}[u(t)] = \Theta^{(\mathrm{R})}[u(-t)]$; this equation is satisfied as long as for at least one time $t$ and one representation the matrix elements of the operators $A(t)$, of the energy and of the state operator $\rho$ are real.

## 2. THE CASE OF LINEAR EQUATIONS OF MOTION

Let us consider a different type of distribution corresponding to the above mentioned ordering principles, for instance, the case of quantum systems subject to linear equations of motion, *i.e.*, systems for which the Lagrangian depends linearly and quadratically on the dynamical variables $q$ [in the case of many discrete degrees of freedom $q = q_1, q_2, \ldots$; in the case of fields $q = \varphi(x)$].

In field theory this case corresponds to free fields, which we assume to be boson-fields. Under those circumstances the commutators of the dynamical variables are $c$-numbers which simplifies matters considerably. The fact is that if repeated commutators of $A$ and $B$ vanish the following formulas hold,

$$e^A e^B = e^B e^{A+[A,\,B]}; \quad e^{A+B} = e^B e^{A+[A,\,B]/2} \tag{10}$$

By applying these formulas several times we are able to find the following correlation between the characteristic functionals $\langle \mathrm{Pr} \exp \int q u dt \rangle$ for the trajectory $q(t)$ corresponding to the principles P, R, $\mathrm{Pr}^{(s)}$,

$$\Theta^{(R)}[u(t)] = \Theta^{(s)}[u(t)]$$
$$= \Theta^{(P)}[u(t)] \exp \frac{1}{2} \int\limits_{t' > t} [q(t), q(t')]\,dt\,dt' \tag{11}$$

Thus for the case of linear equations of motion the symmetrical arrangement gives the same result as the bichronological one. Furthermore, in those cases $q(t)$ is expressed in terms of the initial operators $q(t_0)$, $q'(t_0)$ by means of the same linear transformations as in the classical case. If we substitute this expression into the exponent $\langle \exp \int q u dt \rangle \equiv \Theta^{(s)}[u]$, we see that $\Theta^{(s)}[u]$ can be expressed in terms of the symmetrical initial characteristic function by classical means. Expressed differently, if we start from a Wigner distribution, the time evolution taking into account all multiple distribution laws will not differ from the classical one. As far as the momentaneous distribution in phase space is concerned, this fact was noted in Ref. 1.

Let us consider the correlation function

$$k_{qq}(t, t') = \langle \mathrm{Pr}\, q(t)\, q(t') \rangle - \langle q(t) \rangle \langle q(t') \rangle.$$

Symmetrizing the integral in the exponent (11) we find

$$k_{qq}^{(P)}(t, t') = k_{qq}^{(s)}(t, t') + \frac{1}{2} \operatorname{sgn}(t' - t)\,[q(t'),\, q(t)]. \tag{12}$$

Other correlation functions $k_{(n)\,q}$ $(n = 3, 4, \ldots)$ (higher semi-invariants) are the same for P and $\mathrm{Pr}^{(s)}$.

The results obtained here can directly be generalized to the case of free quantum fields corresponding, for instance, to a Lagrangian $2L = \varphi^\lambda \varphi_\lambda - m^2\varphi^2$. Also we must substitute $\varphi(x)$ for $q(t)$ in (11) and (12) and take into account that

$$\varepsilon(x', x)\,[\varphi(x'), \varphi(x)] = 2i\hbar\bar{\Delta}(x', x) \quad (c = 1).$$

Some initial distributions will give a relativistically invariant correlation function

$$k_{\varphi\varphi}^{(s)}(x, x') = k\Delta^{(1)}(x, x'); \qquad k_{\varphi\varphi}^{(P)} = k\Delta^{(1)} + i\hbar\bar{\Delta}, \tag{13}$$

where $\Delta^{(1)}$ and $\bar{\Delta}$ are the well-known singular functions. If these initial distributions are also normalized we have

$$\Theta^{(s)}[u(x)] = \exp \frac{k}{2} \iint \Delta^{(1)}(x, x')\,u(x)\,u(x')\,dx\,dx'. \tag{14}$$

The ground state corresponds to $k = \hbar/2$ and

$$\hbar^{-1} k_{\varphi\varphi}^{(P)} = \frac{1}{2}\Delta^{(1)} + i\bar{\Delta} = \Delta^F.$$

Having taken the integral inversion one can calculate the distribution densities $w^{(s)}$ and $w^{(P)}$ of various multiplicities, including continuous ones.

The distribution in functional space for the case of the chronological arrangement is closely connected to Feynman's Lagrangian form. We give as an example the probability functional $w^{(P)}[q(t)]$ for a finite time interval $0 \leqslant t \leqslant T$ in the case of a harmonic oscillator having a normalized distribution and a correlation function

$$k_{qq}^{(P)}(t, t') = k\cos\omega(t-t') - \frac{i\hbar}{2m\omega}\sin\omega|t-t'|. \tag{15}$$

The dispersion $k$ of the coordinate fluctuations is connected with $\Theta$ (which is the absolute temperature multiplied by Boltzmann's constant) through the relation $2km\omega = \hbar/\tanh(\hbar\omega/2\Theta)$. Defining

$$\begin{aligned} a &= m\omega/\tanh(\omega T + i\hbar\omega/\Theta); \\ b &= m\omega/\sinh(\omega T + i\hbar\omega/\Theta) \end{aligned} \tag{16}$$

the result of the calculations is in the form

$$\begin{aligned} &w_{0T}^{(r)}[q(t)] \\ &= N^{-1}\exp\frac{i}{\hbar}\left\{\sum_{j=0}^{n-1}\frac{m\varepsilon}{2}\left[\left(\frac{q(t_{j+1})-q(t_j)}{\varepsilon}\right)^2\right.\right. \\ &\left.\left. - \omega^2 q^2(t_j)\right] - \frac{a}{2}[q^2(0)+q^2(T)] + bq(0)q(T)\right\}. \end{aligned} \tag{17}$$

The sum in the exponent in (17) is the action integral $\int_0^{T-\varepsilon} L[q(t+\varepsilon), q(t)]\,dt$. Equation (17) is the Lagrangian form of the thermodynamic Gibbsian quantum ensemble cooresponding to a temperature $\Theta$.

## 3. THE CHRONOLOGICAL ARRANGEMENT IN THE CASE OF ARBITRARY EQUATIONS OF MOTION

When the equations of motion are nonlinear, the commutator $[q(t), q(t')]$ will be a $c$-number only in first order in $t-t'$ (the operator supplement is proportional to $(t-t')^2$). In the nonlinear case, therefore, the results of the previous section are correct for small time intervals to the first order in their duration. This is sufficient to conclude that the Wigner distribution for $q(t)$, $\dot{q}(t)$ follows equally easily from the R as from the $\mathrm{Pr}^{(s)}$ arrangement. However, to determine the distributions involving considerable time intervals we must put our considerations upon a new basis. To begin with we restrict ourselves to the chronological arrangement.

1. We try to determine the density $w^{(P)}(q(t_1), \ldots, q(t_n))$ of the joint distribution of the coordinates at the times $t_1, \ldots, t_n$. If to be definite we put $t_n \geqslant \ldots \geqslant t_1$ we have

$$\begin{aligned} &w^{(P)}(q_1, \ldots . q_n) \\ &= (\Psi^* \delta(q(t_n) - q_n) \ldots \delta(q(t_1) - q_1)\Psi), \end{aligned} \tag{18}$$

where $\Psi$ is the state vector in the Heisenberg representation. We choose as a basis the eigenvectors of the operators $q(t)$. If we go over to a new representation we write Eq. (18) in the form

$$\begin{aligned} w^{(P)}(q_1, \ldots, q_n) = \int\cdots\int \Psi^*(q'_n, t_n)\,\delta(q'_n - q_n)U(q'_n, t_n; q'_{n-1}, t_{n-1})\ldots\delta(q'_1 - q_1)\,\Psi(q'_1, t_1) \\ \times dq'_1 \ldots dq'_n. \end{aligned} \tag{19}$$

where $q'_k = q'(t_k)$, $\Psi(q'(t), t) = (q'(t)|\,)$ is the state vector in the $q(t)$-representation, and $U(q''(t_1), t_1; q'(t), t) = (q''(t_1)|\,q'(t))$ is the unitary operator $U(t_1, t)$ of the transition from one time to another. After integrating we have

$$\begin{aligned} &w^{(P)}(q_1, \ldots, q_n) \\ &= \Psi^*(q_n, t_n)\,U(q_n, t_n; q_{n-1}, t_{n-1})\ldots \\ &\ldots U(q_2, t_2; q_1, t_1)\,\Psi(q_1, t_1). \end{aligned} \tag{20}$$

Dividing the segment $(0, T)$ into increasingly smaller sections by the points $t_k = \varepsilon k$ $(\varepsilon = T/n)$ we can make the transition to a distribution in functional space. As shown by Feynman,[3] we have

$$\begin{aligned} &U(q(t'), t'; q(t), t) = \mathrm{const}\int\ldots \\ &\int\exp\left\{\frac{i}{\hbar}\int_t^{t'-\varepsilon} L[q(t+\varepsilon), q(t)]\,dt\right\}dt\prod_{t+\varepsilon}^{t'-\varepsilon} dq(t), \end{aligned} \tag{21}$$

where $t' > t$, and where $L[q(t+\varepsilon), q(t)]$ is the Lagrangian in which we have replaced $q(t)$ by $[q(t+\varepsilon) - q(t)]/\varepsilon$. By virtue of (20) and (21) the functional probability for a trajectory $q(t)$ in the interval $(0, T)$ will have the form

$$w_{0T}^{(P)}[q(t)] = N^{-1}\Psi^*(q(T),T)\exp\left\{\frac{i}{\hbar}\int_0^{T-\varepsilon} L[q(t+\varepsilon), q(t)]dt\right\}\Psi(q(0),0). \tag{22}$$

The normalizing factor $N$ has a direct meaning only for the final subdivision.

2. The operator (21) can, as is well-known, be written in the form

$$U(q', t'; q, t) = \sum_n \psi_n(q')\exp\left\{-\frac{i}{\hbar}(t'-t)E_n\right\}\psi_n^*(q). \tag{23}$$

where $\psi_n$ is the eigenfunction of the energy operator $H$ corresponding to an eigenvalue $E_n$. We now take in (21) and (23) the limit $t' - t \to \infty$; strictly speaking, the operator $U$ is not defined in that limit. To be able to go to the limit we replace $H$ by $(1-i\mu)H$ where $i\mu$ is small and imaginary. This change corresponds to replacing the Lagrangian $L = T - U$ by

$$L_\mu = (1-i\mu)^{-1}T - (1-i\mu)U \approx L + i\mu H, \tag{24}$$

which expresses the assumption of weak dissipative effects. After this substitution $E_n$ goes over into $(1-i\mu)E_n$ and we have

$$U_\mu(t'-t) \to \psi_0\exp\left\{-\frac{i}{\hbar}(t'-t)E_0\right\}\psi_0^* \quad \text{for} \quad t'-t\to\infty. \tag{25}$$

In agreement with equations (21) and (25) we can write for the wave function of the ground state

$$\psi_0(q(t)) = N^{-1/2}\int\cdots\int\exp\left\{\frac{i}{\hbar}\int_{-\infty}^{t-\varepsilon} L[q(t+\varepsilon), q(t')]dt'\right\}\prod_{-\infty}^{t-\varepsilon} dq(t). \tag{26}$$

where the lower limit $T_1 = -\infty$ must be understood in the following sense: first of all we replace $L$ by $L_\mu$ while $T_1$ is still finite; after that we let $T_1$ go to $-\infty$ simultaneously changing the coefficient $N^{-1/2}$ which depends on $T_1$ and $q(T_1)$; finally we let $\mu \to 0$. In Ref. 6 a different formula was used,

$$\psi_0 = \text{const}\lim_{T\to\infty}\int\cdots\int\exp\left\{\frac{S_T}{\hbar}\right\}\Pi dq(t)\left(S_T = \int_{-T}^{t} L\,dt'\right). \tag{27}$$

If the factor is properly chosen the function $\psi_0$ will be the action function as can be seen, for instance, from Eq. (27). If we take this into account and substitute (26) into (22) we are led to the conclusion that in the case of the ground state the distribution $w_{0T}^{(P)}[q]$ for the trajectory $q(t)$ in the interval $[0, T]$ is equal to the integral of $\exp\frac{i}{\hbar}\int_{-\infty}^{\infty} L\,dt$ over the trajectories outside that interval. Hence it follows naturally that in the case of the ground state the "complete" distribution density for trajectories along the infinite line $-\infty < t < \infty$ is equal to

$$w^{(P)}[q(t)] = N^{-1}\exp\left\{\frac{i}{\hbar}\int_{-\infty}^{\infty} L[q(t)]\,dt\right\}, \tag{28}$$

$$\left(N = \int\cdots\int\exp\left\{\frac{i}{\hbar}\int_{-\infty}^{\infty} L\,dt\right\}\prod_{-\infty}^{\infty} dq(t)\right)$$

This result in its conciseness reminds us of Gibbs' formula for the thermodynamic equilibrium distribution at one definite time. However, Eq. (28) includes a narrower of phenomena, as it refers only to the ground state which in thermo-dynamic language corresponds to the absolute zero. One can go over from the complex distribution (28) to the action functional probability corresponding to the ordering principle (6),

$$w^{(1)}[q] = N_1^{-1}\cos\left\{\frac{1}{\hbar}\int_{-\infty}^{\infty} L[q]\,dt\right\}, \tag{29}$$

## 4. JOINT DISTRIBUTIONS FOR P AND P* TRAJECTORIES

1. If we take into consideration the functional

$$\Theta[u(t), v(t)] = \left\langle P^*\left(\exp\int q(t)v(t)\,dt\right)P\left(\exp\int q(t)u(t)\,dt\right)\right\rangle, \tag{30}$$

we can use it to write for the characteristic functionals for the ordering principles P, P*, and R,

$$\Theta^{(P)}[u] = \Theta[u, 0];\ \Theta^{(P^*)}[u] = \Theta[0, u];$$
$$\Theta^{(R)}[u] = \Theta\left[\frac{u}{2}, \frac{u}{2}\right]. \tag{31}$$

The functional (30) will be interpreted as the characteristic functional $< \exp \int (ru + sv)\, dt >$ of the joint distribution of P-trajectories $r(t)$ and P*-trajectories $s(t)$. If $w[r(t), s(r)]$ is the density of this joint distribution, one can according to equations (31) obtain from it the distribution density of $q(t)$ for the different arrangement principles,

$$w^{(P)}[q(t)] = \int \ldots \int w[q(t), s(t)] \Pi\, ds(t); \tag{31}$$
$$w^{(P^*)}[q(t)] = \int \ldots \int w[r(t), q(t)] \Pi\, dr(t); \tag{32}$$
$$w^{(R)}[q(t)] = \int \ldots \int w\left[q(t) + \frac{\sigma(t)}{2};\ q(t) - \frac{\sigma(t)}{2}\right] \Pi\, d\sigma(t). \tag{33}$$

These formulas express the fact that in order to find the distribution of P-trajectories one must take into account the P*-trajectories and vice versa, and also that the R-trajectories are equal to $(r(t) + s(t))/2$, *i.e.*, are the mean of the P- and the P*-trajectories.

By approximately the same methods as used to obtain (29) and (31) one finds easily for a chosen interval $[0, T]$

$$w_{0T}[r(t), s(t)] = N^{-1} \delta(r(T) - s(T)) \exp\left\{\frac{i}{\hbar} \int_0^{T-\varepsilon} (L[r(t+\varepsilon), r(t)] - L[s(t+\varepsilon), s(t)])\, dt\right\} \psi(r(0), 0)\, \psi^*(s(0), 0). \tag{34}$$

Hence the distributions (32) and (33), and also $w^{(1)}$ are obtained by integration.

2. The formulas found in the previous section are closely connected with the results of Refs. 3 and 7. However, it is impossible to consider the concepts of the present paper to be the same as those of the above mentioned papers. We can illustrate their difference, for instance, by calculating the probability $W$ that in the interval $[0, T]$ the trajectory $q(t)$ shall lie within the limits $a(t) \leqslant q(t) \leqslant b(t)$. In the classical theory this probability is equal to

$$W = \left\langle \prod_0^T \int_{a(t)}^{b(t)} \delta(q(t) - x(t))\, dx(t) \right\rangle. \tag{35}$$

In quantum theory the $q(t)$ are operators and one must in Eq. (35) indicate the order of their arrangement,

$$W = \left\langle \mathrm{Pr} \prod_0^T \int_{a(t)}^{b(t)} \delta(q(t) - x(t))\, dx(t) \right\rangle \tag{36}$$

Different principles Pr give, generally speaking, different probabilities. According to (32) and (33) the principal ones of these are equal to

$$W^{(1)} = \frac{1}{2}\left\{\int_{a(t)}^{b(t)} \ldots \int \int_{-\infty}^{\infty} \ldots \int + \int_{-\infty}^{\infty} \ldots \int \int_{a(t)}^{b(t)} \ldots \int\right\} w[r, s]\, \delta r\, \delta s;$$
$$W^{(2)} = \int_{a(t)}^{b(t)} \ldots \int \int_{-\infty}^{\infty} \ldots \int w\left[q + \frac{\sigma}{2},\ q - \frac{\sigma}{2}\right] \delta q\, \delta\sigma \quad (\delta x = \Pi\, dx(t)). \tag{37}$$

These quantities differ in the limits of integration from the quantity

$$W = \int_{a(t)}^{b(t)} \ldots \int \int_{a(t)}^{b(t)} \ldots \int w[r, s]\, \delta r\, \delta s, \tag{38}$$

which was indicated in Refs. 3 and 7 and which represents the absolute square of the probability of the event considered. In our opinion expression (38) cannot be interpreted as a probability since there exists no arrangement principle ordering Pr such that expression(38) would follow from equation (36). Expression (38) is thus not connected with any complete statistical picture.

## 5. STATISTICAL INTERPRETATION BASED UPON THE R-ARRANGEMENT

1. Processes of time evolution described by the distributions $w^{(P)}[q(t)]$, $w^{(P^*)}[q(t)]$ are Markov processes, *i.e.*, the joint distribution density of the values $q(t_1), \ldots, q(t_n)$ can be written in the form

$$w(q(t_1), \ldots, q(t_n)) = V(q(t_n), t_n; q(t_{n-1}), t_{n-1}) \ldots \ldots V(q(t_2), t_2; q(t_1), t_1)\, W(q(t_1), t_1). \tag{39}$$

Indeed, Eq. (39) agrees with Eq. (20) if we put

$$W(q, t) = \Psi^*(q, t)\,\Psi(q, t);\quad V(q', t'; q; t) = \frac{\Psi^*(q', t')}{\Psi^*(q, t)}\,U(q', t'; q, t).$$

Unfortunately, the Markov transition probability $V$ turns out to depend on the chosen state $\Psi$. The arrangements $\mathrm{Pr}^{(1)}$ and R do not lead to Markov processes, even in such a sense. However, the R arrangement corresponds to a "Markov process with derivatives", *i.e.*, the transition from a pair of values $q(t)$, $\dot{q}(t)$ to the same pair at a different time is a Markov process. Expressed differently, evolution in phase space is a genuine Markov process (with a transition probability independent of the state). To see this we consider the joint R-distribution of the quantities $q(t_1)$, $q(t_1 + \varepsilon)$, $q(t_2)$, $q(t_2 + \varepsilon), \ldots, q(t_n + \varepsilon)$ which we obtain by writing expression (33) in a discrete version instead of in a continuous one. It has the form

$$\begin{gathered}\int\ldots\int U(q'_n, q_n + \tau_n)\,U^*(q'_n, q_n - \tau_n)\,U(q_n + \tau_n, q'_{n-1} + \tau'_{n-1})\\ \times U^*(q_n - \tau_n, q'_{n-1} - \tau'_{n-1})\,U(q'_{n-1} + \tau'_{n-1}, q_{n-1} + \tau_{n-1})\\ \times U^*(q'_{n-1} - \tau'_{n-1}, q_{n-1} - \tau_{n-1})\ldots U(q_2 + \tau_2, q'_1 + \tau'_1) \times U^*(q_2 - \tau_2, q'_1 - \tau'_1)\\ \times U(q'_1 + \tau'_1, q_1 + \tau_1)\,U^*(q'_1 - \tau'_1, q_1 - \tau_1)\,R(q_1 + \tau_1, q_1 - \tau_1)\\ 2^{2n-1}\,d\tau_1\ldots d\tau_n\,d\tau'_1\ldots d\tau'_{n-1},\end{gathered} \tag{40}$$

where $q_k = q(t_k)$, $q'_k = q(t_k + \varepsilon)$, $\tau_k = \sigma_k/2$, $R$ is the density matrix, $U(q_k, q'_{k-1}) = U(q(t_k), t_k; q(t_{k-1} + \varepsilon), t_{k-1} + \varepsilon)$ is the operator (21).

Taking it into account that for small values of $\varepsilon$

$$U(q'_k, q_k) \approx c\,\exp\left\{\frac{i}{\hbar}\,\frac{m}{2\varepsilon}\,(q'_k - q_k)^2\right\},\quad c = \left(\frac{im}{2\pi\hbar\varepsilon}\right)^{1/2}, \tag{41}$$

we write Eq. (40) in the form

$$\begin{gathered}|c|^{2n}\int\ldots\int \exp\left\{-\frac{im}{\hbar\varepsilon}\,(q'_n - q_n)\,\sigma_n\right\} U\left(q_n + \frac{\sigma_n}{2},\ q'_{n-1} + \frac{\sigma'_{n-1}}{2}\right)\\ \times U^*\left(q_n - \frac{\sigma_n}{2},\ q'_{n-1} - \frac{\sigma'_{n-1}}{2}\right)\exp\left\{\frac{im}{\hbar\varepsilon}\,(q'_{n-1} - q_{n-1})\,(\sigma'_{n-1} - \sigma_{n-1})\right\}\ldots\\ \ldots U\left(q_2 + \frac{\sigma_2}{2},\ q'_1 + \frac{\sigma'_1}{2}\right)U^*\left(q_2 - \frac{\sigma_2}{2},\ q'_1 - \frac{\sigma'_1}{2}\right)\\ \times\exp\left\{\frac{im}{\hbar\varepsilon}\,(q'_1 - q_1)\,(\sigma'_1 - \sigma_1)\right\}R\left(q_1 + \frac{\sigma_1}{2},\ q_1 - \frac{\sigma_1}{2}\right)d\sigma_1\ldots d\sigma_n\,d\sigma'_1\ldots d\sigma'_{n-1}.\end{gathered} \tag{42}$$

The last expression splits into a product of functions. Since the Jacobian of the transformation from

$$q(t_1),\ q(t_1 + \varepsilon), \ldots, q(t_n),\ q(t_n + \varepsilon)$$

to the quantities

$$q(t_1), \dot{q}(t_1), \ldots, q(t_n), \dot{q}(t_n)\quad \left(\dot{q}(t) = \frac{q(t + \varepsilon) - q(t)}{\varepsilon}\right)$$

is equal to $\varepsilon^{-n}$ the density of the distribution of the latter quantities differs from expression (42) by that factor and is equal to

$$w^{(R)}(q(t_1), \dot{q}(t_1), \ldots, \dot{q}(t_n)) = K_{t_n t_{n-1}}(q_n, \dot{q}_n; q_{n-1}, \dot{q}_{n-1})\ldots K_{t_2 t_1}(q_2, \dot{q}_2; q_1, \dot{q}_1)\,W_t(q_1, \dot{q}_1), \tag{43}$$

where

$$W_t(q, \dot{q}) = \frac{m}{2\pi\hbar}\int \exp\left\{-\frac{im}{\hbar}\dot{q}\sigma\right\} R\left(q+\frac{\sigma}{2},\ q-\frac{\sigma}{2}\right)d\sigma; \tag{44}$$

$$K_{t't}(q', \dot{q}'; q, \dot{q}) = \frac{m}{2\pi\hbar}\iint \exp\left\{-\frac{im}{\hbar}(\dot{q}'\sigma' - \dot{q}\sigma)\right\} U\left(q'+\frac{\sigma'}{2},\ q+\frac{\sigma}{2}\right)U^*\left(q'-\frac{\sigma'}{2},\ q-\frac{\sigma}{2}\right)d\sigma\, d\sigma' \tag{45}$$

are the corresponding Wigner distribution and Markov transition probability. Thus the evolution process is actually a Markov process in the case of the R arrangement.

The function (45) was considered by Moyal[1] as the operator defining the time evolution of the momentaneous distribution law (44). In particular, he derived a differential equation for it. We note that Moyal's statement that this evolution process is a Markov process is in his paper only a hypothesis inasmuch as the multiple distribution laws do not follow from the evolution of the momentaneous distribution law. As we have just shown, this hypothesis is justified for the R-arrangement of operators, but just for the $\Pr^{(s)}$ arrangement, which was used by Moyal as a basis to derive the Wigner distribution, it is not justified.

2. In many cases the integral (33) can be obtained from expression (34). A number of these cases correspond to quantum fields with localized interactions described by the usual form of the interaction Lagrangian. We consider as an example the relativistic case of the interaction of a neutral with a charged field. For the sake of simplicity we assume both fields to be scalar. Let the Lagrangian be of the form

$$L[\varphi, \psi] = \frac{1}{2}(\varphi_\lambda\varphi^\lambda - \mu^2\varphi^2) + \psi^*_\lambda\psi^\lambda - m^2\psi^*\psi + g\varphi\psi^*\psi. \tag{46}$$

The generalization of equations (33) and (34) for the case of the distribution of these fields within a four-dimensional region $M$ is given by the equation

$$\begin{aligned} w^{(R)}[\varphi(x), \psi(x)] = N^{-1}\int\ldots\int \exp\left\{\frac{i}{\hbar}\int_M dx\left(L\left[\varphi+\frac{\xi}{2},\ \psi+\frac{\eta}{2}\right]\right.\right. \\ \left.\left. - L\left[\varphi-\frac{\xi}{2},\ \psi-\frac{\eta}{2}\right]\right)\right\}\prod_M d\xi(x)\, d\eta(x)\, d\eta^*(x). \end{aligned} \tag{47}$$

We have in this equation not written out explicitly the boundaries of the region $M$. From Eq. (46) we get

$$\begin{aligned} F(x) \equiv L\left[\varphi+\frac{\xi}{2},\ \psi+\frac{\eta}{2}\right] - L\left[\varphi-\frac{\xi}{2},\ \psi-\frac{\eta}{2}\right] = \varphi^\lambda\xi_\lambda - \mu^2\varphi\xi + g\psi^*\psi\xi \\ + \eta^*_\lambda\psi^\lambda - m^2\eta^*\psi + g\varphi\eta^*\psi + \psi^{*\lambda}\eta_\lambda - m^2\psi^*\eta + g\psi^*\eta + \frac{1}{4}g\xi\eta^*\eta. \end{aligned} \tag{48}$$

We subdivide the region $M$ into cells of size $\Delta_j$, containing the points $x_j$. We replace the integral in the exponent in expression (47) by the sum $\Sigma F(x_j)\Delta_j$, where $F(x_j)$ is defined by expression (48) in which the derivatives $\xi_\lambda$, $\eta_\lambda$ are represented by the differences of the values of the $\xi(x_j)$, $\eta(x_j)$ in neighboring cells. After substituting this expression into the sum we combine all those terms in the exponent which contain the same variables $\xi(x_j)$, $\eta(x_j)$, and we get

$$\int\ldots\int \exp\left\{\frac{i}{\hbar}\sum_j F_j\Delta_j\right\}\prod_j d\xi_j\, d\eta_j\, d\eta^*_j = \prod_j\int \exp i\{\alpha_j\xi_j + \beta^*_j\eta_j + \beta_j\eta^*_j + \gamma_j\varepsilon_j\eta^*_j\eta_j\}d\xi_j\, d\eta_j\, d\eta^*_j, \tag{49}$$

where

$$\alpha_j = \frac{\Delta_j}{\hbar}[\Box\varphi - \mu^2\varphi + g\psi^*\psi]_{x=x_j};\ \beta_j = \frac{\Delta_j}{\hbar}[(\Box - m^2 + g\varphi)\psi]_{x=x_j};$$

$$\gamma_j = \Delta_j/4h;\ \xi_j = \xi(x_j),\ \eta_j = \eta(x_j).$$

The factors on the right hand side of (49) refer to those cells which are not near the boundary of the region considered. We have not written out explicitly the factors referring to the cells on the boundaries as they are more complicated. Integration over $\eta_j$ and $\eta_j^*$ in (49) means integrating both the real and the imaginary part from $-\infty$ to $+\infty$. Let $\eta_j = \rho + i\sigma$, $\beta_j = \rho_1 + i\sigma_1$; one can take each factor of the integral separately and we obtain

$$\int \exp i\{\alpha\xi + 2(\rho\rho_1 + \sigma\sigma_1) + \gamma\xi(\rho^2 + \sigma^2)\}d\xi\,d\rho\,d\sigma = \frac{2\pi^2}{\gamma} I_0\left(2\left[\frac{\alpha}{\gamma}(\rho_1^2 + \sigma_1^2)\right]^{1/2}\right), \tag{50}$$

where $I_0(x) = J_0(ix)$ is a Bessel function of the second kind. We have thus

$$w^{(R)}[\varphi(x), \psi(x)] = N_0^{-1}\prod_j I_0(2(\alpha_j\beta_j^*\beta_j/\gamma_j)^{1/2})$$

$$= N_1^{-1}\left[\exp\left\{\frac{1}{2\hbar\sqrt{g}}\int[\Box\varphi - \mu^2\varphi + g\psi^*\psi]^{1/2}\,|\,(\Box - m^2 + g\varphi)\psi\,|\,z(x)\,dx\right\}\right]^{(2)}. \tag{51}$$

In this equation$^{(2)}$ indicates a symbolic square defined by the equation

$$\left[\sum_{\lambda,\mu,\ldots} a_{\lambda\mu\ldots} z_1^\lambda z_2^\mu \ldots\right]^{(2)} = \sum_{\lambda,\mu,\ldots}(a_{\lambda\mu\ldots})^2, \tag{52}$$

so that

$$[e^{az}]^{(2)} = I_0(2a);\ [e^{\Sigma a_j z_j}]^2 = \Pi I_0(2a_j).$$

If we write out in detail the factors in Eq. (51) referring to the interior points of the region we get in the limit $\hbar\sqrt{g} \to \infty$ the "classical" form

$$\prod_j \delta(\Box\varphi - \mu^2\varphi + g\psi^*\psi)\,\delta((\Box - m^2 + g\varphi)\psi)\,\delta \quad \text{(compl. conj.)},$$

$$(\delta(\alpha + i\beta)\,\delta(\alpha - i\beta) \equiv \delta(\alpha)\,\delta(\beta)).$$

It is well known that one obtains the "classical" basis of evolution not only as $\hbar \to 0$, but also in the case of linear equations of motion (Sec. 2).

## 6. CONCLUDING REMARKS

We shall make several remarks about the connection between the statistical interpretation just developed and well-known methods of relativistic field theory.

1. Let $A(x)$ stand for all field operators (for instance, $\varphi$, $\psi$) and let $L[A]$ be the corresponding Lagrangian. For each arrangement principle there is for a given complete statistical picture a corresponding characteristic functional $\Theta = \langle S\rangle$ where

$$S = \Pr\exp\left\{i\int A(x)\,J(x)\,dx\right\}.$$

We remember that in the classical case [commuting $A(x)$] we have for every $F[A]$

$$F\left[\frac{\delta}{i\delta J(x)}\right]S = F[A(x)]\,S. \tag{53}$$

In quantum theory the operation of $F[\delta/i\delta J(x)]$ on $S$ gives $\Pr\{F[A]\,S\}$ which, generally speaking, is different from $F[A]\,S$. We shall get a formula very similar to the classical expression (53) in the case of the P (and P*) arrangement. In this case after "disentangling" from $\mathrm{P}\{F[A]S\}$ the exponential factor[4] we shall have

$$F\left[\delta/i\delta J(x)\right] S = F\left[\tilde{A}(x)\right] S, \tag{54}$$

where $\tilde{A}(x)$ are the operators "in a different representation," more precisely the operators whose evolution is described by the Lagrangian $L[A] + AJ$ (we put $\hbar = 1$).

If one considers that the addition of $AJ$ describes external sources, the fields $A(x)$, which evolve without sources, will be operators in the "interaction representation" (with respect to the "interaction" $AJ$). Also $S$ will be the "scattering matrix" describing the action of external sources. The characteristic functional $\Theta$ in the P-arrangement can thus be interpreted dynamically as the probability amplitude for a $\Psi - \Psi$ transition under the action of external sources.

If $F_1[A(x)] = F_2[J(x)]$ is the equation of motion of the fields with sources, then in virtue of (54) we can go over to the following equation for the characteristic functional describing fields without sources

$$F_1\left[\delta / i\delta J(x)\right]\Theta = F_2\left[J(x)\right]\Theta.$$

Equations of this form were used in Ref. 8 to derive Feynman's Lagrangian form and in Ref. 9 to derive Schwinger's equations. Our aim was to draw attention to the fact that $\Theta$ ($Z$ in the notation of the above mentioned papers) can equally well be treated as a characteristic functional.

2. We define generalized moment and correlation (cumulative) functions by the equations

$$m_n(x_1, \ldots, x_n \mid J) = \frac{1}{\Theta}\frac{\delta}{i\delta J(x_1)} \cdots \frac{\delta}{i\delta J(x_n)}\Theta; \tag{55}$$

$$k_n(x_1, \ldots, x_n \mid J) = \frac{\delta}{i\delta J(x_1)} \cdots \frac{\delta}{i\delta J(x_n)} \ln\Theta. \tag{56}$$

The usual moment and correlation functions are defined analogously, but after performing the differentiations one puts $J(x) \equiv 0$. The difference between the functions (55) and (56) for $J \neq 0$ and the usual ones is essentially that the probability functional $w[A]$ which describes the field distribution, is replaced by a new functional

$$\exp\left(i\int AJ\,dx\right) w[A] \Big/ \int \exp\left(i\int AJ\,dx\right) w[A]\,\delta A,$$

for which expressions (55) and (56) are the usual moment and correlation functions. The formulas in which expressions (55) and (56) occur will thus have the usual form.

In view of this all we see that the distribution functions introduced by Schwinger[10] for the general case when external sources are present are nothing but the correlation functions (56), and some equations involving then can be explained in a trivial manner. For instance, a formula typical for Ref. 10 of the form

$$\frac{\delta\langle A(x')\rangle}{i\delta J(x)} = \frac{\delta^2 \ln\Theta}{i\delta J(x)\, i\delta J(x')} = \langle PA(x)\, A(x')\rangle - \langle A(x)\rangle\langle A(x')\rangle$$

is essentially an elementary formula from correlation theory

$$k_2(x, x') = m_2(x, x') - m_1(x)\, m_1(x').$$

3. Let the Lagrangian $L$ be the sum of the Lagrangian of free fields $L_0$ and an interaction Lagrangian $L_J$. The ground state $\Phi_0$ of the free fields and the ground state $\Psi_0$ of the interaction field correspond to different P-distributions (28), namely

$$w_0^{(P)}[A] = N_0^{-1}\exp\left\{i\int_{-\infty}^{\infty} L_0\,dx\right\};$$
$$w^{(P)}[A] = N^{-1}\exp\left\{i\int_{-\infty}^{\infty} L\,dx\right\}. \tag{57}$$

Hence we have

$$\int F[A]\, w\,\delta A = \frac{N_0}{N}\int F[A]\exp\left\{i\int L_J\,dx\right\} w_0\,\delta A. \tag{58}$$

Integrating over all $w$ means averaging over the state $\Psi_0$ and over all $w_0$ averaging over the state $\Phi$. Taking this into account we can go over to the quantum mechanical average, and if we furthermore compare expression (58) with itself, but with $F = 1$, and use the result, we get

$$(\Psi_0^* P F(A)\,\Psi_0) = (\Phi_0^* P(F[A]\,S)\,\Phi_0)/S_0;$$
$$S_0 = N/N_0$$
$$\equiv \int \exp\left(i\int L\,dx\right)\delta A \Big/ \int \exp\left(i\int L_0 dx\right)\delta A. \tag{59}$$

In this equation $S_0 = (\Phi_0^* S \Phi_0)$, and $S$ is the scattering matrix produced by the interaction $L_J$. Thus the use of the distribution (57) makes it easier to derive the relations connected with the interaction representation.

We note that from equation (26) applied to $\Psi_0$ and $\Phi_0$ we can derive the relation $\Psi_0 = N_0^{1/2} N^{-1/2} S(0, -\infty)\,\Phi_0$, which forms the contents of Gell-Mann and Low's theorem;[11] in our opinion it is necessary to intro-

duce into $L_0$ and $L$ an infinitely small dissipation in order to prove this theorem (Sec. 3).

[1] J. E. Moyal, Proc. Camb. Phil. Soc., 45, 99 (1949).
[2] R. L. Stratonovich, J. Exptl. Theoret. Phys. (U.S.S.R.) 31, 1012 (1956); Soviet Phys. JETP 4, 891 (1957).
[3] R. P. Feynman, Revs. Mod. Phys. 20, 467 (1948).
[4] R. P. Feynman, Phys. Rev. 84, 108 (1951).
[5] P. A. M. Dirac, Revs. Mod. Phys. 17, 195 (1945).
[6] R. P. Feynman, Phys. Rev. 97, 660 (1955).
[7] I. M. Gel'fand and A. M. Iaglom, Usp. Mat. Nauk 11, 77 (1956); translation scheduled for publication in Fortschr. d. Phys.
[8] I. M. Gel'fand and R. A. Minlos, Dokl. Akad. Nauk (U.S.S.R.) 97, 209 (1954).
[9] E. S. Fradkin, J. Explt. Theoret. Phys. (U.S.S.R.) 29, 121 (1955); Soviet Phys. JETP 2, 148 (1956).
[10] J. Schwinger, Proc. Nat. Acad. Sci. 37, 452 (1951).
[11] M. Gell-Mann and F. Low, Phys. Rev. 84, 350 (1951).

Translated by D. ter Haar
291

SOVIET PHYSICS JETP VOLUME 5, NUMBER 6 DECEMBER 15, 1957

# The Surface Impedance of Metals in the Infrared Region

M. I. KAGANOV AND V. V. SLEZOV
*Physico-Technical Institute, Ukranian Academy of Sciences*

(Submitted to JETP editor November 2, 1956)
J. Exptl. Theoret. Phys. (U.S.S.R.) 32, 1496-1504 (June, 1957)

An expression has been obtained for the surface impedance of metals in the infrared region without making any special assumptions concerning the law of dispersion of the conduction electrons.

**1.** IN THE OPTICS OF METALS the most interesting region, it seems to us, is the infrared region in which the frequency of the electromagnetic field satisfies the condition

$$\nu_0 \ll \omega \ll \omega_a \qquad (1)$$

Here $\nu_0 = 1/\tau$ is the collision frequency ($\tau$ is the relaxation time), and $\omega_a$ is the limit of the internal photoeffect.* As shown in Ref. 1, for low temperatures and for pure metals this region is known to be essentially: $\nu_0 \sim 10^{11}$, $\omega_a \sim 10^{15}$ to $10^{16}$. Thus we talk of working at wavelengths on the order of ten microns.

In this frequency region the electron gas is approximately described by the dielectric constant

$$\varepsilon = 1 - 4\pi Ne^2/m\omega^2 \qquad (2)$$

where $N$ is the density of free electrons and $m$ the effective mass.* In other words, the electron gas "behaves" like an electron plasma.

Since $\Omega^2 \approx 4\pi Ne^2/m \sim \omega_a^2$, then in the frequency region of interest to us the dielectric constant of the metal is negative ($\varepsilon < 0$), and, its modulus is considerably greater than 1, *i.e.*,

$$\varepsilon \approx -4\pi Ne^2/m\omega^2, \quad |\varepsilon| \gg 1.$$

The reflection of light from the surface of a metal in this case is, in principle, not connected with the ohmic loss and can be described by a purely imaginary surface impedance

* The internal photoeffect which is due to the interaction of electrons can occur at all frequencies; however, it principally occurs only for $\omega > \omega_a$ ($\hbar\omega_a$ is of the order of intervals between energy bands, $\sim 10^{-12}$ ergs).

* Note that according to the opinion of Ginzburg (*cf.* Ref. 1) we put in formula (2) the mass of the free electrons. All resultant changes in $N/m$ are due to the change in the electron density $N$. This, of course, cannot lead to contradiction, but, it seems to us, causes trouble in comparing optically measured constants of the electron gas with results of other experiments (galvanomagnetic, specific heat, *etc.*). Moreover, if the calculation is carried out with the use of the kinetic equation, it is natural that in formula (2) there should appear precisely the effective mass (see below).