STOCHASTIC PROCESSES IN PSEUDOEUCLIDEAN SPACE

L. V. PROKHOROV

Leningrad State University

Submitted to JETP editor June 18, 1966

J. Exptl. Theoret. Phys. (U.S.S.R.) 52, 167–174 (January, 1967)

A model is considered in which a point is subjected to random displacements in a pseudoeuclidean plane. This model does not allow for a description in terms of a probability density only (i.e., within the framework of usual probability theory). An attempt to enlarge the mathematical formalism leads to the conclusion that it is possible to find a probabilistic theory in which a solution to the proposed problem exists and which turns out to be unique under sufficiently broad assumptions, and is such that the mathematical formalism and the postulates are identical to those of quantum mechanics.

1. We consider the following model: a point is subjected to random displacements in a pseudoeuclidean plane. In addition we require the following conditions to be satisfied: i) invariance with respect to pseudoeuclidean rotations; ii) independence of the displacements along the two axes and equivalence of the two directions along each of the axes.

We wish to find the probability density for finding the point in a given region of the plane. We denote this quantity by w and consider that, in addition to its dependence on the initial and final coordinates of the point, it also depends on an invariant, nondecreasing parameter τ which we agree to call the "proper time" of the wandering point. We further assume homogeneity of the two-dimensional space and of the τ axis. Thus the problem becomes completely defined. But before starting to analyze it, we wish to stress the fact that the model is a purely mathematical construction, and at this stage no attempt is being made to attach any physical interpretation to it.

Thus, if at the "instant" τ_1 (we shall omit the quotation marks in the sequel) the point had the coordinates x_1, the probability density for finding it at the point x_2 at the instant τ_2 can depend only on the differences of the corresponding variables: $w(x_2 - x_1, \tau_2 - \tau_1)$, owing to the assumed homogeneity of the space and the variable τ.

Further, it follows from assumption i) that the process under consideration must not single out any direction in space. This implies that the function $w(x, \tau)$ can depend only on the invariants which can be formed out of the vector x, namely on x^2 and $\theta(x_0)\theta(x^2)$. The condition ii) allows the function to depend only on x^2, and by the definition of the function $\tilde{w}(x^2, \tau) \equiv w(x, \tau)$, the following condition must be satisfied

$$\tilde{w}_1(x_0^2, \tau)\tilde{w}_1(-x_1^2, \tau) = \tilde{w}(x_0^2 - x_1^2, \tau), \qquad x^2 \equiv x_0^2 - x_1^2 \tag{1}$$

(independence of the displacements along the two axes). This condition leads to the following explicit form of the function $w(x, \tau)$:

$$w(x, \tau) = A(\tau)e^{\alpha(\tau)x^2}, \tag{2}$$

where $A(\tau)$ and $\alpha(\tau)$ are functions of the parameter τ. However, the function (2) cannot be interpreted as a probability density for any function $\alpha(\tau)$ since both for $\alpha(\tau) > 0$ and for $\alpha(\tau) < 0$ the normalization integral diverges:

$$\int w(x, \tau)d^2x \equiv A(\tau)\int e^{\alpha(\tau)(x_0^2 - x_1^2)}dx_0\,dx_1 = \infty. \tag{3}$$

(In fact, even condition i) implies that w is not normalizable, but we shall make use of the form (2) which was determined by making use of the condition (ii).)

We have thus arrived at a contradiction. The question arises how to interpret the contradiction we have found. Does (3) mean that the process is impossible under the conditions i) and ii) we have imposed? Or is it a consequence of the mathematical formalism we have used? Apparently it is more correct to assume the second point of view, since we shall show below that a probabilistic scheme exists such that the proposed model admits a noncontradictory description within its framework.

2. We try to use a more flexible mathematical machinery, assuming that the model is based not on a probability density $w(x, \tau)$ but on an auxil-

iary quantity $v(x, \tau)$ which allows a unique determination of $w(x, \tau)$. Conditions i) and ii) will be imposed on this new quantity $v(x, \tau)$, rather than on the probability density. The function $v(x, \tau)$ does not need to be assumed to be either positive or even real. Repeating the reasoning that led to Eq. (2) (in particular, reinterpreting the condition ii) as the validity of (1) for the new function $v(x, \tau)$) we obtain a similar expression for v:

$$v(x,\tau) = A(\tau)\exp\{\alpha(\tau)x^2\}. \tag{4}$$

where $A(\tau)$ and $\alpha(\tau)$ are complex-valued functions[1].

By specifying a relation between the auxiliary quantity $v(x, \tau)$, which we shall call probability amplitude, and the probability density $w(x, \tau)$, we determine the structure of the probabilistic scheme. We restrict our attention to the "local" case, when $w(x, \tau)$ depends on the values of $v(x, \tau)$ only at the same point, and is independent of the values taken at other points. This means that the relation between w and v is of the form

$$w(x,\tau) = \varphi(v(x,\tau)), \tag{5}$$

where $\varphi(v)$ is a nonnegative function.

We now find the form of this function. Since the applicability of Eq. (1) was assumed for the amplitudes $v(x, \tau)$, this means that the amplitude of a composite event must be the product of the amplitudes of the independent events. The same must also be true for probability densities. This leads to the following functional equation for the function φ:

$$\varphi(v_1 v_2) = \varphi(v_1)\varphi(v_2), \tag{6}$$

which has as positive solutions the functions of the form

$$\varphi(v) = |v|^p, \tag{7}$$

with p a real number. Taking into account the normalization condition for w, we conclude that the probability amplitudes must belong to the function space L_p and must be normalized:

$$\int |v(x,\tau)|^p\, d^2x = 1. \tag{8}$$

We now turn to the explicit form (4) of the function $v(x, \tau)$ which we have determined before. We note first that the function $\alpha(\tau)$ must be regarded as pure imaginary:

$$\operatorname{Re}\alpha(\tau) = 0 \tag{9}$$

(otherwise the probability density turns out to be exponentially increasing at infinity). But then $|v(x, \tau)| = |A(\tau)|$, i.e., independent of x, and therefore $w(x, \tau)$ is again not normalizable. It seems that we have accomplished nothing. However the crux of the matter is the fact that the probability amplitude (4) is the transition amplitude from an initial distribution which is not normalizable, therefore the distribution at time τ it is also not normalizable. But before showing this we make the form of the function $v(x, \tau)$ more precise by imposing an additional condition: iii) the probability amplitude $v(x, \tau)$ is subjected to the composition law (convolution):

$$v(x,\tau) = \int v(x-x', \tau-\tau')\, v(x', \tau')\, d^2x'. \tag{10}$$

The requirement (10) for the amplitudes is new and independent. But since we propose as a starting point the amplitudes and impose physical requirements on these quantities, rather than on the $w(x, \tau)$ as is usually done, the requirement (10) looks quite natural. In describing Brownian motion in Euclidean space this condition is imposed on $w(x, \tau)$ and is a consequence of the axioms of the theory of probability.[2]

It is easy to see that the condition (10) leads to the following relations for $\alpha(\tau)$ and $A(\tau)$ (in relation to the function $v(x, \tau)$ represented by Eq. (4), with $\operatorname{Re}\alpha(t) = 0$, the integral (10) is an improper integral, but can easily be made meaningful):

$$\frac{1}{\alpha(\tau-\tau')} + \frac{1}{\alpha(\tau')} = \frac{1}{\alpha(\tau)}, \tag{11}$$

$$A(\tau-\tau')A(\tau')\frac{\pi}{\alpha(\tau-\tau')+\alpha(\tau')} = A(\tau). \tag{12}$$

The solution of these functional equations is given by

$$\alpha(\tau) = \frac{k}{\tau}, \quad A(\tau) = \frac{k}{\pi\tau}e^{k_1\tau}. \tag{13}$$

In deriving Eqs. (11)—(13) we have substituted $\alpha(\tau) \to i\alpha(\tau)$, in order to deal with a real function $\alpha(\tau)$. Therefore the arbitrary constant k in (13) is real whereas the other arbitrary constant k_1 must always be chosen pure imaginary, in order to avoid exponential growth (or decay) of the probability as $\tau \to \infty$. For simplicity we set this constant equal to zero for the time being. This

[1]In general, nothing compels us to choose the functions $A(\tau)$ and $\alpha(\tau)$ as complex-valued functions. One could have chosen for $\alpha(\tau)$ an antihermitian matrix. There may exist schemes with an even more complicated mathematical structure.

[2]In place of iii) one could postulate the superposition principle for amplitudes (cf. Sec. 3, condition b)). Then Eq. (10) would be a consequence.

will not affect the fundamental results of the present section. Thus, finally we derive the following expression for the amplitude $v(x, \tau)$:

$$v(x,\tau) = \frac{k}{\pi\tau} e^{ikx^2/\tau}. \tag{14}$$

We now prove that if at $\tau = 0$ we have a normalized distribution with amplitude $f(x)$, i.e.,

$$\int |f(x)|^p d^2x = 1, \tag{15}$$

there exists a unique p for which this property is maintained at an arbitrary instant τ, i.e., such that

$$\int |f(x,\tau)|^p d^2x = 1, \tag{16}$$

where, according to conditions ii) and iii), the function $f(x, \tau)$ is given by the equation

$$f(x,\tau) = \int v(x - x', \tau) f(x') d^2x', \tag{17}$$

where $v(x, \tau)$ is given by (14). Indeed, choosing the concrete initial distribution

$$f(x) = (ap/\pi)^{1/p} \exp\{-a(x_0^2 + x_1^2)\}, \quad a = \text{const} > 0,$$

which is normalized according to (15), it is easy to compute $f(x, \tau)$. The absolute value of this function turns out to have the following expression:

$$|f(x,\tau)| = \frac{k}{b\tau}\left(\frac{ap}{\pi}\right)^{1/p} \exp\left\{-\frac{k^2}{\tau^2}\frac{a}{b^2}(x_0^2 + x_1^2)\right\}, \tag{18}$$

with $b^2 = a^2 + k^2/\tau^2$. The norm of $f(x, \tau)$ is determined by

$$\int |f(x,\tau)|^p d^2x = \left(\frac{k}{b\tau}\right)^{p-2}, \tag{19}$$

which shows that the normalization is conserved only for a theory with $p = 2$. Thus a formalism of the type of quantum mechanics is the only acceptable one.

It remains to verify that for $p = 2$ the conservation of norm holds for an arbitrary initial distribution. But this follows simply from the remark that (14) satisfies the equation

$$\int v^*(x - x_1, \tau) v(x - x_2, \tau) d^2x = \delta(x_1 - x_2). \tag{20}$$

Here $\delta(x)$ is the two-dimensional delta-function.

It becomes clear now that the choice of the function (14) as probability amplitude at time τ means that at time $\tau = 0$ the initial amplitude $f(x)$ was a delta-function $\delta(x)$, i.e., not square-integrable. For the same reason the norm of $v(x, \tau)$ turned out to be infinite. In contradistinction, the probability density $w(x, \tau)$, even if we disregard its non-normalizability, transforms an initially normalized distribution into an unnormalizable one. This is easy to see by noting that the delta-function is a normalizable distribution in this case.

In conclusion we discuss the fact that the amplitude $f(x, \tau)$ we have obtained is still a noninvariant function (in addition to the vector x it also depends some supplementary constant vectors, as a consequence of its normalizability). But this noninvariance (the appearance of additional vectors) is not related to the properties of the space or to peculiarities of the random displacements of the point, but is exclusively related to the noninvariant character of the initial distribution, which cannot be invariant since it is normalizable. In quantum mechanics a similar noninvariance would be related to the presence of macroscopic bodies which participate in the preparation of the system (formation of the initial distribution) but which do not affect the ultimate development of the motion.

3. We formulate explicitly the rules of operation with probability amplitudes, which are consequences of the arguments of the preceding section. There are two such rules:

3. We formulate explicitly the rules of operation with probability amplitudes, which are consequences of the arguments of the preceding section. There are two such rules:

a) The probability amplitude for a composite event is the product of the amplitudes of independent events (applicability of Eq. (1) to amplitudes).

b) The probability amplitude for the realization of at least one of a set of incompatible events is the sum of the amplitudes of the individual events (superposition principle for probability amplitudes; it follows from an interpretation of Eq. (10) for amplitudes). In an axiomatic approach to probability these two rules, together with the rule for calculation of probabilities from a known probability amplitude, should be included among the axioms characterizing a given probabilistic scheme.

There remains one important question to be discussed. The rules a) and b) are also true for the probabilities themselves and therefore it is important to indicate under what circumstances they should be applied to probability amplitudes and when the probability densities. In quantum mechanics the choice between these two possibilities is dictated by the measuring process. If a system is not subjected to measurements one should use the rules a) and b) for amplitudes. A measurement process implies a transition from amplitudes to probabilities. Therefore one has to introduce into the formalism under consideration a concept similar to the concept of quantum-mechanical measurement.

Let the point go from the state a into the state c via the intermediate state b. The amplitude for this transition is given by the formula

$$\psi_{ac} = \sum_b \psi_{ab}\psi_{bc}, \qquad (21)$$

where ψ_{ab} and ψ_{bc} are respectively the transition amplitudes from a to b, and from b to c. If one does not attempt to specify the model further, one can formally describe the "measuring process" in the intermediate state b as the operation which transforms the function $|\sum_b \psi_{ab}\psi_{bc}|^2$ into $\sum_b |\psi_{ab}|^2 |\psi_{bc}|^2$. In doing this, one should generally treat any transition from amplitudes to probabilities as a result of a "measuring process." This makes the formalism well defined.

After introducing complex amplitudes the model automatically exhibits other quantum-mechanical traits also: the existence of unitary-equivalent representations, hermitian operators corresponding to "observables," uncertainty relations, etc. This is all easily established and we shall not do it in detail (cf. the Appendix). We only remark the following noteworthy fact which is characteristic for the model under consideration. Returning to Eq. (13) we consider the general case when $k_1 \neq 0$ and set $k_1 = -im/2$. We also set $k = m/2$, i.e., choose $v(x, \tau)$ of the form

$$v(x,\tau) = \frac{m}{2\pi\tau} \exp\left\{-i\frac{m}{2}\left(\frac{x^2}{\tau} + \tau\right)\right\}. \qquad (22)$$

If one assumes that there is no way of "measuring" the quantity τ, i.e., that from the localization of the point in a region one cannot infer anything about the value of τ, we should in fact be interested in the probability amplitude for finding the point in a given region for any τ. But in order to obtain such an amplitude it is necessary to integrate the function (22) over τ from 0 to ∞ (superposition over τ). As a result of this integration one obtains the Stueckelberg-Feynman propagator for a particle of mass m (up to a normalization constant which has to be introduced at the same time as the τ-integration).

Indeed, the integration over τ of the Fourier transform of the amplitude (22) leads to the relation

$$\frac{2mi}{p^2 - m^2 + i\varepsilon} = \int_0^\infty \exp\left\{i\frac{\tau}{2m}(p^2 - m^2 + i\varepsilon)\right\} d\tau, \qquad (23)$$

where one should take the limit $\epsilon \to 0$. The assumptions underlying the model under consideration make it impossible to attribute a serious meaning to this formula. At the same time it is not completely devoid of interest, since the model process described by it could be considered as some relativistically invariant process occurring in a two-dimensional space[3], where the role of time is played by the coordinate x_0.

4. To summarize, we have established the following. First, it was shown that there exist stochastic processes which cannot be described by means of the usual concept of probability theory only, but which can be described in terms of probability amplitudes.

There are, of course, models for which both the usual and the quantum-mechanical descriptions do not lead to contradictions (for example Brownian motion in Euclidean space; in this case the description in terms of probability amplitudes coincides with the description given in terms of nonrelativistic quantum mechanics). The applicability of one or another of the schemes depends in this case on the nature of the model under consideration.

Secondly, under sufficiently broad assumptions, the quantum-mechanical scheme (description in terms of probability amplitudes belonging to the Hilbert space L_2) turned out to be the only possible one for the model under consideration. This means that in a certain sense one could take about the "uniqueness" of the quantum-mechanical description.

Another fact is worth mentioning, namely the fact that only "classical" objects occurred in the model: the mathematical point and the pseudo-euclidean space. Finally, in connection with the fact that in quantum mechanics motion is described by means of a probability amplitude, there arises the question as to the physical causes which are responsible for the necessity of just such a description, i.e., the causes of the quantum-mechanical behavior of elementary particles. In the model we have discussed the cause was the pseudo-euclidean character of the space of random displacements of the point.

Another interesting question is whether there are other schemes, differing from quantum mechanics, but allowing a probabilistic interpretation.

The author would like to express his profound gratitude to Yu. V. Novozhilov, V. A. Fock and V. A. Franke for an interesting discussion and valuable remarks.

APPENDIX

We show how the basic elements of the mathematical formalism of quantum mechanics appear in the scheme under consideration.

---

[3]A similar formula holds for Minkowski space also.

Operators and commutation relations. The expectation value of the coordinate of the point is given by

$$\bar{x}_\mu(\tau) = \int x_\mu w(x,\tau)\, d^2x = \int f^*(x,\tau)\, x_\mu f(x,\tau)\, d^2x. \quad (A.1)$$

The expectation value of the velocity is defined by

$$\bar{v}_\mu = \frac{d}{d\tau}\bar{x}_\mu(\tau) = \int \left[\frac{\partial f^*}{\partial \tau} x_\mu f + f^* x_\mu \frac{\partial f}{\partial \tau} \right] d^2x. \quad (A.2)$$

Let $f(x, \tau)$ be defined by (17), where $v(x, \tau)$ is chosen as (22), i.e., we consider the general situation with $k_1 \neq 0$. It is easy to check that for $\tau > 0$ the amplitude $f(x, \tau)$ satisfies the equation

$$2mi\frac{\partial f(x,\tau)}{\partial \tau} + (\Box - m^2) f(x,\tau) = 0, \quad (A.3)$$

where $\Box \equiv -\partial^2/\partial x_0^2 + \partial^2/\partial x_1^2$. This equation is identical to Fock's equation with proper time for the free particle[1]. Substituting into Eq. (A.2) the expression for the derivative $\partial f/\partial \tau$ given by (A.3), an integration by parts leads to the following expression for $\bar{v}_\mu$:

$$\bar{v}_\mu = \frac{i}{m} g_{\mu\nu} \int f^*(x,\tau) \frac{\partial}{\partial x_\nu} f(x,\tau)\, d^2x. \quad (A.4)$$

Here $g_{\mu\nu}$ is the metric tensor: $g_{00} = 1$, $g_{11} = -1$. It follows from this equation that the expectation value of the velocity of the point equals the expectation value of the operator $(i/m) g_\mu \partial/\partial x_\nu$, therefore it is natural to associate this operator to the velocity of the point:

$$\hat{v}_\mu = \frac{i}{m} g_{\mu\nu} \frac{\partial}{\partial x_\nu}. \quad (A.5)$$

This implies the commutation relations for $\hat{v}_\mu$ and $\hat{x}_\nu$:

$$[\hat{v}_\mu, \hat{x}_\nu] = \frac{i}{m} g_{\mu\nu}. \quad (A.6)$$

Transformation theory. The function $f(x, \tau)$ determines the probability density for finding the point with a coordinate x. We transform to a new function by means of the unitary operator $U(y, x)$

$$\tilde{f}(y,\tau) = \int U(y,x) f(x,\tau)\, d^2x. \quad (A.7)$$

The unitarity of this operator implies that $\tilde{f}(y, \tau)$ is also normalized and consequently $|\tilde{f}(y, \tau)|^2$ can be interpreted as the probability density describing the distribution of y. An example of such a unitary transformation is a Fourier transform. Since this transformation maps the differentiation operator into multiplication by a number, if we remember that the velocity operator is represented by a differentiation, it is easy to see that the Fourier transform of the function $f(x, \tau)$ is the probability amplitude for the velocity distribution of the point.

Uncertainty relations. Since the velocity operator (A.5) and the position operator $\hat{x}$ are subject to the commutation relations (A.6), the uncertainty relations can be derived by means of a standard procedure (for instance by means of Weyl's trick[2]). These relations are of the form

$$\overline{(\Delta v_\mu)^2}\,\overline{(\Delta x_\mu)^2} \geqslant \frac{1}{4m^2} \quad (A.8)$$

(no summation over μ).

[1] V. A. Fock, Izv. AN SSSR, ser. fiz. No. 4—5, 551 (1937).

[2] L. D. Landau and E. M. Lifshitz, Kvantovaya Mekhanika (Quantum Mechanics) Fizmatgiz, 1963 (2nd Russian Edition), p. 69.

Translated by M. E. Mayer
22