Œuvres de Laurent Schwartz

Laurent Schwartz

**Density of probability of presence of elementary particles**

*Proc. 4th Berkeley Sympos. Math. Statist. and Prob., vol. III*, Berkeley, Calif.: Univ. California Press, 1961, p. 307-314.

Extrait des *Œuvres de Laurent Schwartz*
publiées par la Société mathématique de France, 2011.

# DENSITY OF PROBABILITY OF PRESENCE OF ELEMENTARY PARTICLES

LAURENT SCHWARTZ
UNIVERSITY OF PARIS

## 1. Introduction: nonrelativistic case

In the initial, nonrelativistic theory of quantum mechanics it is assumed that the only information we have about the state of a particle, at a given time, is its wave function $\Psi$, a complex function on $R^3$ or a complex function of three coordinates $x$, $y$, $z$. This function is assumed to be square integrable, $\Psi \in L^2$, and moreover one assumes

$$\iiint_{R^3} |\Psi(x, y, z)|^2 \, dx \, dy \, dz = 1. \tag{1.1}$$

Consider an observable physical quantity, taking its values in a set $X$. For example, the position of the particle is a quantity with values in $X = R^3$ and so is the velocity. The energy has values in $X = R$, and so on. In classical mechanics, a measurement of such a quantity is supposed to be obtainable with arbitrary accuracy, and, for a given state, the quantity has a definite value $x$ in $X$. In quantum mechanics, this unlimited precision disappears. If we make a measurement of the quantity, for a particle having the wave function $\Psi$, we have only a probability law $P_\Psi$, depending on $\Psi$, that is, a positive measure on $X$, of total mass 1. Thus, if $A$ is a subset of $X$, assumed to be measurable $(P_\Psi)$, the probability that the measurement will give a result in $A \subset X$ is $P_\Psi(A)$. It is usually assumed that this probability law $P_\Psi$ on $X$ must be given by a spectral decomposition of the Hilbert space $L^2$, with respect to $X$. Such a spectral decomposition is defined as follows. It is a map $P: A \to P(A) = L^2_A$, where $A$ runs over a Borel field of subsets of $X$, and $L^2_A$ is a closed subspace of $L^2$, with the following properties.

(a) $L^2_\phi = \{0\}$, where $\phi$ = empty set of $X$, $0$ = origin of the vector space $L^2$; $L^2_X = L^2$.

(b) If $A$ and $B$ are disjoint subsets of $X$, $L^2_A$ and $L^2_B$ are orthogonal in $L^2$.

(c) If $A$ is the union of a finite or denumerable family of disjoint subsets $A_n$, then $L^2_A$ is the closure of the subspace of $L^2$ spanned by the $L^2_{A_n}$.

Thus the probability law $P_\Psi$ of the physical quantity under consideration must be given by

$$P_\Psi(A) = ||\Psi_A||^2 = \iiint\limits_{R^3} |\Psi_A(x, y, z)|^2 \, dx \, dy \, dz, \tag{1.2}$$

where $\Psi_A$ is the orthogonal projection of $\Psi \in L^2$ on the subspace $P(A) = L^2_A$ of $L^2$. Axiom (a) ensures that $P_\Psi(\phi) = 0$, $P_\Psi(X) = 1$, and (b) and (c) ensure, according to Pythagoras' theorem, that $P_\Psi$ is a completely additive set function; it is therefore, as desired, a probability law on $X$.

In this model the state of the particle is given by the wave function $\Psi$, the observable physical quantity by the spectral decomposition $P$, and thus a measurement of the quantity for the state of the particle is governed by the probability law $P_\Psi$ on $X$, given by (1.2). There is in $L^2(R^3)$ a trivial spectral decomposition, that for which $L^2_A$ is the subspace of those $\Psi$ which are zero outside $A$. It is regarded as the spectral decomposition associated with the observable: "position of the particle in $R^3$." Therefore, we have, in a measurement, the following probability for the particle to be found in the subset $A$ of $R^3$

$$P_\Psi(A) = \iiint\limits_{A} |\Psi(x, y, z)|^2 \, dx \, dy \, dz. \tag{1.3}$$

For this reason, $|\Psi|^2$ is the density of probability of presence. If now we look for the spectral decomposition corresponding to the first coordinate $x$ of the particle, taking its values in $R$, it must be that for which, when $B \subset R$, the probability for a measurement of $x$ to give a result in $B$ is

$$\iiint\limits_{x \in B} |\Psi(x, y, z)|^2 \, dx \, dy \, dz. \tag{1.4}$$

This spectral decomposition is also the spectral decomposition associated with the self-adjoint operator on $L^2$ "multiplication by $x$." Multiplication by $x$ is also said to be the operator associated with the measurement of $x$.

If now we consider the evolution in time of the given particle, we shall have, at every instant $t$, a wave function $\Psi_t$, and thus a function of time having values in $L^2$. It will also be a function $\Psi$ of the four variables $x, y, z, t$, defining for every $t$ a function $\Psi_t$ of the three variables $x, y, z$. The usual rules of quantum mechanics say that $\Psi$ must satisfy some Schrödinger equation such as

$$i \hbar \frac{\partial \Psi}{\partial t} = H_t \Psi_t, \tag{1.5}$$

where, for every $t$, the $H_t$ is a self-adjoint operator on $L^2$. This self-adjointness ensures, according to known properties of Hilbert spaces, that any solution of (1.5) keeps the same norm in $L^2$ for every $t$; if, for $t = 0$, it has the norm 1, which is required by (1.1), this equality remains valid for every $t$, and the solution defines a valid wave function for every $t$, and finally a valid motion of the particle. The Hamiltonian $H$, or the function $t \to H_t$, depends on the mechanical conditions under consideration.

## 2. Relativistic case

In special relativity no distinction is made between the three space variables $x$, $y$, $z$ and the time variable $t$. The universe is a space $E_4$, a four-dimensional *affine space*, having an associated *vector space* $\vec{E}_4$. We note that $E_4$ is not a vector space, it has no origin, and there is no sum of any two points; $\vec{E}_4$ is the space of vectors of $E_4$. If $a$ and $b$ are two points of $E_4$, then $\overrightarrow{b - a}$ is a vector, belonging to $\vec{E}_4$. On $E_4$ is given a quadratic form, with signature (3.1) measuring the "universe lengths." A physical coordinate system is an orthonormal basis of $E_4$, given by an origin of $E_4$ and four vectors of $\vec{E}_4$. If $x_1, x_2, x_3, x_4 = ct$ are called the corresponding coordinates of an event (an *event* is a point of $E_4$), the observer sees $x_1$, $x_2$, $x_3$ as its space coordinates and $t$ as its time. A sense of time and an orientation of $\vec{E}_4$ are also given.

The complete motion of a particle will be a wave function $\Psi$, a complex function on $E_4$. For a physical coordinate system $\Psi$ becomes a function of four variables $x_1$, $x_2$, $x_3$, $t$, and we are led back to the situation of section 1.

We shall consider that a given particle in given mechanical conditions is characterized by all its possible motions. We may assume that all these possible motions will be all the elements of norm 1 in a Hilbert space $\mathcal{H}$ of functions on $E_4$. For instance, in the nonrelativistic case, for a particle characterized by the Hamiltonian $H$, the Hilbert space $\mathcal{H}$ was formed by all the functions $\Psi$ of four variables $x$, $y$, $z$, $t$ satisfying (1.5) and belonging to $L^2_{x,y,z}$ for every $t$. The norm in $\mathcal{H}$ was given by

$$(2.1) \qquad ||\Psi||^2_{\mathcal{H}} = \iiint_{R^3} |\Psi(x, y, z, t)|^2 \, dx \, dy \, dz,$$

the result being independent of $t$ because of the self-adjointness of $H_t$. We can certainly not have the same kind of results in the relativistic case, because it is not Lorentz invariant.

It is an uninteresting restriction to force $\Psi$ to be a function; we shall only assume $\Psi$ to be a distribution on $E_4$, a wave distribution. Remember that a distribution $\Psi$ is a continuous linear form on the space $\mathcal{D}(E_4)$ of the infinitely differentiable functions on $E_4$, with compact support. The value of $\Psi$ on $\varphi \in \mathcal{D}$ will be denoted by $\Psi(\varphi)$ or $\langle \Psi, \varphi \rangle$. $\mathcal{H}$ will be a subspace of the space $\mathcal{D}'(E_4)$ of the distributions on $E_4$. $\mathcal{H}$ will also have a given structure as a Hilbert space, and we shall assume that the norm in $\mathcal{H}$ is such that convergence in $\mathcal{H}$ implies convergence in the sense of distributions. There are infinitely many choices of $\mathcal{H}$, each of which gives a possible particle in some well-defined physical situation, and all the $\Psi \in \mathcal{H}$, with norm 1 in $\mathcal{H}$, give all the possible motions of such a particle in the situation considered. We are only interested in spaces $\mathcal{H} \neq \{0\}$ since we have to deal with elements of $\mathcal{H}$ of norm 1.

## 3. Free scalar elementary particle

For a detailed proof of the formulas given here, such as (3.3) and (4.5), see Schwartz [1].

If the particle is free (no external fields), it has to be Lorentz universal, or Lorentz invariant, in the sense that a Lorentz transformation on a possible motion $\Psi$ must give a new possible motion.

Thus we shall assume that, for any element $\sigma$ of the Lorentz group and any $\Psi$ of $\mathcal{H}$, the transformed distribution $\sigma\Psi$ also belongs to $\mathcal{H}$, and has the same norm in $\mathcal{H}$, that is,

$$||\sigma\Psi||_{\mathcal{H}} = ||\Psi||_{\mathcal{H}}. \tag{3.1}$$

Note that the transformed distribution $\sigma\Psi$ is defined, for any function $\varphi \in \mathcal{D}$ which is infinitely differentiable with compact support, by

$$\sigma\Psi(\varphi) = \Psi(\sigma^{-1}\varphi) = \Psi[\varphi(\sigma x)]. \tag{3.2}$$

Therefore, $\sigma$ is a unitary operator on the Hilbert space $\mathcal{H}$, and the Lorentz group $G$ has here a unitary representation in $\mathcal{H}$. We shall define as a *free elementary particle* a free particle (thus Lorentz invariant) for which $\mathcal{H}$ is *minimal*, in the sense that no Lorentz invariant Hilbert space $\mathcal{H}' \neq \{0\}$ contained in $\mathcal{H}$ exists except $\mathcal{H}' = \mathcal{H}$ with a proportional norm. Therefore the unitary representation of the Lorentz group $G$ is simply an irreducible unitary representation.

What we call here the Lorentz group is the *proper inhomogeneous Lorentz group*, that is, the group of all the affine operators of $E_4$ onto itself, preserving the given quadratic form, on $\vec{E}_4$, the orientation, and the sense of time. The word *inhomogeneous* simply means that we consider affine operators of $E_4$ (for example, translations), and the word *proper* means that we restrict ourselves to operators preserving orientation (determinant $+1$) and sense of time.

The complete list of all these Hilbert spaces $\mathcal{H} \subset \mathcal{D}'(E_4)$, Lorentz invariant and minimal, may be obtained by different techniques, all using Fourier transforms. The result is the following. Of course, for every $\mathcal{H}$, one can also take the same with a proportional norm, but we shall not distinguish them.

(a) There is one special $\mathcal{H}$, one-dimensional, all the elements of which are constant functions $\Psi$. It may be interpreted as the vacuum.

(b) There is a series of spaces $\mathcal{H}_1$, depending on one real parameter. These cannot be physically interpreted.

(c) There is a normal series, physically interpretable. It depends on a parameter $m_0 \geqq 0$, which may be interpreted as the rest mass of the particle, and a parameter $\pm$, which may be interpreted as the electric charge.

In this way the only particles we have found are the $\pi$-mesons, with spin 0. We find, in this way, every possible mass $m_0$, including 0, which is not true in nature! One can generalize and find all the known elementary particles by looking for finite-dimensional vector-valued elementary particles, for which $\Psi$ is finite-dimensional vector-valued, that is, $\Psi$ has a finite number of scalar components. We find here charged particles only, because we considered complex-valued wave distributions $\Psi$. With real-valued distributions neutral particles are obtained.

The Hilbert space $\mathcal{H}_{m_0,+}$ may be described in the following way. Consider the distribution on $\vec{E}_4$

$$(3.3)\qquad 2\pi\Delta^{-}_{\frac{2\pi c m_0}{h}}(\vec{X})$$

$$= \text{p.v.}\left[\frac{\frac{\pi c m_0}{2h} N_1\left(\frac{2\pi c m_0}{h}\sqrt{-\vec{X}^2}\right)}{\sqrt{-\vec{X}^2}} Y(-\vec{X}^2) + \frac{\frac{c m_0}{h} K_1\left(\frac{2\pi c m_0}{h}\sqrt{\vec{X}^2}\right)}{\sqrt{\vec{X}^2}} Y(X^2)\right]$$
$$+ i\left[\frac{\frac{\pi c m_0}{h}\epsilon(X_0) J_1\left(\frac{2\pi c m_0}{h}\sqrt{-\vec{X}^2}\right)}{\sqrt{-\vec{X}^2}} Y(-\vec{X}^2) - \frac{1}{2}\epsilon(X_0)\delta(\vec{X}^2)\right].$$

In this rather complicated formula $N_1$ is a Neumann function; $K_1$ a Kelvin function; $J_1$ a Bessel function (one could use a shorter formula with Hankel functions); $\Delta^-$ is the name of the distribution, one of the "singular functions," that is, distribution of quantum mechanics; p.v. means Cauchy's principal value; $\vec{X}^2$ means the value on the vector $\vec{X} \in \vec{E}_4$ of the Lorentz quadratic form; $Y$ means the Heaviside function where $Y(\tau) = 1$ for $\tau \geqq 0$, and $= 0$ for $\tau < 0$; $\epsilon$ is defined as the function $\epsilon(\tau) =$ sign of $\tau = +1$ for $\tau \geqq 0$, and $-1$ for $\tau < 0$, so that if $X_0$ is the fourth component of $\vec{X}$ in any coordinate system $X_1, X_2, X_3, X_0$ then $\epsilon(X_0)$, for elements $\vec{X}$ of the interior or the surface of the light cone, is $+1$ for $\vec{X}$ in the positive light cone, $-1$ for $X$ in the negative light cone; $\delta(\vec{X}^2)$ is defined from the $\delta(u)$ of one variable $u$ by the change of variables $u = \vec{X}^2$ (we denote here distributions in the physical way, as functions); $c$ is the velocity of light; $h$ is Planck's constant. The parameter is written $cm_0/h$ so that $m_0$ may be interpreted as rest mass of the particle. Then a distribution $\Psi$ on $E_4$ belongs to $\mathcal{H}_{m_0],+}$ if and only if the expression

$$(3.4)\qquad \frac{|\langle\Psi, \varphi\rangle|}{\langle 2\pi\Delta^{-}_{\frac{2\pi c m_0}{h}} * \bar{\varphi}, \varphi\rangle^{1/2}}$$

is bounded when $\varphi$ runs over $\mathfrak{D}(E_4)$. Here $*$ means convolution. In this case the upper bound is the norm of $\Psi$ in $\mathcal{H}_{m_0,+}$.

All the $\Psi$ of $\mathcal{H}_{m_0,+}$ are solutions of the Klein-Gordon equation

$$(3.5)\qquad \Box\Psi - \frac{4\pi^2 c^2 m_0^2}{h^2}\Psi = 0.$$

This equation is here *not assumed;* we find it as a *consequence* of our hypothesis that $\mathcal{H}$ is Lorentz invariant and minimal.

The Hilbert space $\mathcal{H}_{m_0,-}$ is obtained in the same way from $\Delta^+$, which is obtained from $\Delta^-$ by changing $i$ into $-i$.

## 4. Density of probability of presence

From now on we shall write $\mathcal{H}$ instead of $\mathcal{H}_{m_0,\pm}$. Then every $\Psi$ of $\mathcal{H}$ is a priori a distribution. Actually one can prove it is a function, that is, a locally integrable function defined almost everywhere on $E_4$.

Consider a physical coordinate system. Thus $\Psi$ becomes a function of $(x, y, z, t)$, locally integrable, defined almost everywhere. Therefore, if we fix the time $t = t_0$, then $\Psi$ is *not defined* as a function of $x, y, z$, since a hyperplane $t = t_0$ is a set of measure zero in $E_4$. But one can prove the following result: it is possible to choose $\Psi$ (initially defined only almost everywhere) so that it is a *continuous function* of $t$, with values in the space $L^1_{\text{loc}}$ of the locally integrable functions of $(x, y, z)$. Because of the continuity in $t$, the function $\Psi$ is then determined not merely almost everywhere in $E_4$ but, *for every* $t$, almost everywhere with respect to $(x, y, z)$.

Finally, $\Psi$ defines for $t = t_0$, a well-defined Lebesgue class of functions $\Psi_{t_0}$, and also a well-defined distribution $\Psi_{t_0}$ on $R^3$. Moreover, it can be proved that a knowledge of $\Psi_{t_0}$, *the cross section of* $\Psi$ *over the hyperplane* $t = t_0$, completely determines $\Psi$ (quantum-mechanical determinism). The system of the function $\Psi_{t_0}$ is a subspace $\mathcal{H}_{t_0}$ of $\mathcal{D}'(R^3)$, having a one-to-one correspondence $\Psi \to \Psi_{t_0}$ with $\mathcal{H}$. Carrying over the Hilbert structure of $\mathcal{H}$ onto $\mathcal{H}_{t_0}$, we define $\mathcal{H}_{t_0}$ as a Hilbert space contained in $\mathcal{D}'(R^3)$, which may be called the cross section of the Hilbert space $\mathcal{H}$ by $t = t_0$. Now any physical observable quantity at the time $t_0$, with values in a set $X$, must be measured by a spectral decomposition of $\mathcal{H}_{t_0}$, relative to $X$. If $A \to P(A) = (\mathcal{H}_{t_0})_A$ is this spectral decomposition, the probability of finding the value of a measurement of the quantity in $A$, when the wave function is $\Psi \in \mathcal{H}$, with $||\Psi|| = 1$, will be

$$P_\Psi(A) = ||(\Psi_{t_0})_A||^2, \tag{4.1}$$

where $(\Psi_{t_0})_A$ is the orthogonal projection of $\Psi_{t_0}$ on $(\mathcal{H}_{t_0})_A$. Since $||\Psi_{t_0}|| = ||\Psi||$, where the Hilbert structure on $\mathcal{H}_{t_0}$ is defined by carrying over that of $\mathcal{H}$, we have that $P_\Psi$ is, as desired, a probability law on $X$. We are interested in the measurement of the position of the particle at the time $t_0$, whose physical quantity, the position, has values in $R^3$. Here the result is essentially different from that of the nonrelativistic case. One cannot postulate that the manifold $(\mathcal{H}_{t_0})_A$ is formed by all the $\Psi$ equal to zero outside $A$, because, as is seen by studying the scalar product in $\mathcal{H}_{t_0}$, in this case $(\mathcal{H}_{t_0})_A$ and $(\mathcal{H}_{t_0})_B$ would not be orthogonal subspaces in $\mathcal{H}_{t_0}$. In other words, $|\Psi_{t_0}|^2$ cannot be the density of probability of presence. In yet other words, the "position operator" in coordinate $x_i$ for $i = 1, 2, 3$, cannot be multiplication by $x_i$, as it is in the nonrelativistic case, because such an operator is not self-adjoint in the Hilbert space $\mathcal{H}_{t_0}$. In the physical literature a density of probability of presence for the meson is often considered which is not even positive! What should be the spectral decomposition relative to $R^3$, corresponding to the measurement of position at the time $t_0$?

It is natural to ask whether there exists a one-to-one norm preserving, linear transformation $\Psi_{t_0} \to \oplus$, from $\mathcal{H}_{t_0}$ onto $L^2(R^3)$, that is, covariant with the inhomogeneous proper orthogonal group $\Gamma$ of $R^3$. That is, $\Psi_{t_0}$ must be such that

$$\Psi_{t_0} \to \oplus \text{ implies } \tau\Psi_{t_0} \to \tau\oplus \tag{4.2}$$

whenever $\tau$ belongs to $\Gamma$. Note that $\Gamma$ is the group of affine operators of $R^3$, preserving lengths and the orientation. Here inhomogeneous means that it contains the translations, proper that it preserves the orientation.

In this case the trivial spectral decomposition of $L^2(R^3)$ will define a spectral decomposition of $\mathcal{H}_{t_0}$ and $(\mathcal{H}_{t_0})_A$ will be the set of $\mathcal{H}_{t_0}$ corresponding to the set of $L^2$ formed by all the $\oplus$ equal to zero outside $A$. Such a spectral decomposition will be acceptable as a spectral decomposition for the measurement of the position of the particle at the time $t_0$, and $|\oplus|^2$ will be acceptable as a possible density of probability of presence at the time $t_0$, for the particle having the wave function $\Psi$ or the instantaneous $t_0$-wave function $\Psi_{t_0}$.

In fact, such a map $\Psi_{t_0} \to \oplus$ can be found. It is given as follows. If $\Delta$ is a Laplacian on $R^3$, by Fourier transform $\mathfrak{F}$ there is classically defined an operator

$$\sqrt{2}\left(-\frac{\Delta}{4\pi^2} + \frac{c^2m_0^2}{h^2}\right)^{1/4}. \tag{4.3}$$

Thus, one has

$$\oplus = \sqrt{2}\left(-\frac{\Delta}{4\pi^2} + \frac{c^2m_0^2}{h^2}\right)^{1/4}\Psi \tag{4.4}$$

or

$$\mathfrak{F}\oplus = \sqrt{2}\left(\rho^2 + \frac{c^2m_0^2}{h^2}\right)^{1/4}\mathfrak{F}\Psi_{t_0}, \tag{4.5}$$

where $\rho$ is the distance from the origin.

Actually, it may be written as a convolution,

$$\oplus = \Psi_{t_0} * \frac{\left[2^6\left(\frac{cm_0}{h}\right)^7\pi^{-1}\right]^{1/4}}{\Gamma\left(-\frac{1}{4}\right)}\rho^{-7/4}K_{7/4}\left(2\pi\frac{cm_0}{h}\rho\right), \tag{4.6}$$

where $K$ is a Kelvin function, decreasing, classically, exponentially at infinity. As we observe, $\oplus$ is obtained from $\Psi_{t_0}$ by a convolution, which is a *nonlocal* operation. Therefore, knowledge of $\Psi_{t_0}$ in an open set $\Omega$ of $R^3$ does not allow us to know $\oplus$ in $\Omega$; for this, a complete knowledge of $\Psi_{t_0}$ is necessary.

There are infinitely many other isometries of $\mathcal{H}_{t_0}$ onto $L^2(R^3)$ having the same property of covariance with the orthogonal group. Namely, one can take the previous one followed by any unitary transformation of $L^2$ onto itself, commuting with the inhomogeneous proper orthogonal group of $R^3$. Such a unitary transformation $\oplus' \to \oplus''$ is given, using the Fourier transform $\mathfrak{F}$, by the formula

$$\mathfrak{F}\oplus'' = (\mathfrak{F}\oplus')e^{if(\rho)}, \tag{4.7}$$

where $f(\rho)$ is an arbitrary measurable function of the distance $\rho$ from the origin of $R^3$.

Therefore, all the possible operations $\Psi_{t_0} \to \oplus$ are of the form

$$\oplus = \mathfrak{L} * \Psi_{t_0}, \tag{4.8}$$

where

$$\mathfrak{F}\mathfrak{L} = \sqrt{2}\left(\rho^2 + \frac{c^2 m_0^2}{h^2}\right)^{1/4} e^{if(\rho)}. \tag{4.9}$$

Since the coefficient of $e^{if(\rho)}$ is real and nonnegative, while $e^{if(\rho)}$ itself is never real and nonnegative unless $f(\rho) = 2k\pi$, it can be seen that there is one and only one transformation of the form (4.9), where $\mathfrak{L}$ is a distribution of positive type, having a positive measure as Fourier transform. But I do not see any physical reason for $\mathfrak{L}$ to be of positive type.

I should rather think that in the correspondence between the physical particle and the mathematical representation, there remains some arbitrariness. One example is the choice of $\mathfrak{L}$, and the simplest choice is given in (4.6). The same can be done for vector-valued (spin) particles.

REMARK. Of course, the formulas and equations given here are well known in physics; only the point of view and the method of exposition are new (and, eventually, the mathematical rigor!).

Our density of probability of presence was already introduced by Newton and Wigner [2].

## REFERENCES

[1] L. SCHWARTZ, *Matemática y Física Cuántica*, Buenos Aires, Universidad de Buenos Aires, 1958.

[2] T. D. NEWTON and E. P. WIGNER, "Invariant theoretical determination of position operators," *Phys. Rev.*, Vol. 76 (1949), p. 191(A).