Relativistic probability amplitudes I. Massive particles of any spin

Scott E. Hoffmann
School of Mathematics and Physics,
The University of Queensland,
Brisbane, QLD 4072
Australia[*]

We consider a massive particle of arbitrary spin and the basis vectors that carry the unitary, irreducible representations of the Poincaré group. From the complex coefficients in normalizable superpositions of these basis vectors, we identify momentum/spin-component probability amplitudes with the same interpretation as in the nonrelativistic theory. We find the relativistic transformations of these amplitudes, which are unitary in that they preserve the modulus-squared of scalar products from frame to frame. Space inversion and time reversal are also treated. We reconsider the Newton-Wigner construction of eigenvectors of position and the position operator. Position/spin-component probability amplitudes are also identified and their relativistic, unitary, transformations derived. Again, space inversion and time reversal are considered. For reference, we show how to construct positive energy solutions of the Klein-Gordon and Dirac equations in terms of probability amplitudes. We find the boost transformation of the position operator in the spinless case and present some results on the relativity of position measurements. We consider issues surrounding the classical concept of causality as it applies in quantum mechanics. We briefly examine the relevance of the results presented here for theories of interaction.

[*] scott.hoffmann@uqconnect.edu.au

I. INTRODUCTION

This paper will deal entirely with free particles, or particles isolated far from other particles with which they might interact. While we will not construct a relativistic theory of interaction here, we will see that an understanding of the behaviour of free particles is an essential addition to any such theory.

Historically, the inclusion of the principles of relativity into quantum mechanics has met with serious problems, with the Klein-Gordon equation as our first example [1, 2]. There, it was assumed that a spinless particle should be described by an amplitude that transforms relativistically as a scalar function, $\varphi(x)$, of position and time ($x^\mu = (t, \boldsymbol{x})^\mu$). A manifestly covariant equation (in terms of the four-vector derivative $\partial_\mu = \partial/\partial x^\mu$) was proposed to determine the time evolution. This is the Klein-Gordon equation for mass m_0,

$$\{\partial^\mu \partial_\mu + m_0^2\}\varphi(x) = 0. \tag{1}$$

Problems arose immediately. The equation is second order in the time derivative, unlike the nonrelativistic Schrödinger equation, changing the nature of the initial value problem. The interpretation of the amplitude, $\varphi(x)$, is unclear. It was evident that its modulus-squared, $|\varphi(x)|^2$, could not be interpreted as a probability density, as in the nonrelativistic theory. It does not obey an invariant conservation equation. Then a locally conserved four-vector current density, $J_{\text{KG}}^\mu(x) = i\{\varphi^*(x)\partial^\mu\varphi(x) - \partial^\mu\varphi^*(x)\varphi(x)\}$ was constructed, to be proposed as a probability current density. But the zero component is not positive definite.

In addition, the Klein-Gordon equation has unphysical negative energy solutions as well as positive energy solutions. This should be obvious, since the equation is expressing "$p^2 = m_0^2$" instead of "$p^0 = +\sqrt{\boldsymbol{p}^2 + m_0^2}$". A particle with a negative energy has a velocity in the opposite direction to its momentum. Allowing the inclusion of negative as well as positive energies in the one superposition leads to rapid, real oscillations of both momentum and position amplitudes, destroying any possibility of normalization.

The four-component Dirac amplitude [1, 2] for spin-1/2 electrons and positrons also has an interpretation problem. The index of the four components is not the eigenvalue of an observable and the values of the components change when the gamma matrix representation is changed. The triumph of Dirac's theory, though, is that it constructs a locally transforming four-vector current density with negative/positive definite zero component for electrons/positrons. The four-component amplitude is perhaps best seen as a necessary intermediate step in this construction. As with the Klein-Gordon equation, the Dirac equation allows unphysical negative energy solutions.

None of these problems is present in the theory given below.

An alternative approach, reviewed here and expanded upon, begins with the basis vectors carrying the unitary, irreducible representations of the Poincaré group (the group of spacetime translations and Lorentz transformations). These basis vectors are constructed to have only positive energies, in accordance with physical observations. In fact the operator representing time reversal is antiunitary precisely so that the positivity of energy is preserved under this transformation. Normalized superpositions of these basis vectors can be formed to describe wavepacket states. The relativistic transformations of the superposition amplitudes are to be derived rather than postulated. We come to realize that relativity allows these, generally more complicated, transformation properties (we call this non-manifest covariance) as well as the manifest covariance of scalars, four-vectors and tensors.

From these complex superposition coefficients we can identify relativistic momentum/spin-z-component probability amplitudes for a massive particle of any spin. Then we refer to the construction by Newton and Wigner [3] of position/spin-z-component eigenvectors for a massive particle with any spin. This will lead us to identify relativistic position/spin-z-component probability amplitudes. We derive their transformation properties, which are found to be, for boosts, nonlocal.

In Section II we derive the relativistic transformation properties of the momentum/spin-z-component probability amplitudes for a massive particle with general spin. Many of these results have already been explained by Fong and Rowe [4]. However it appears that it is still not widely known that probability amplitudes taking the same roles of those in the nonrelativistic theory can be consistent with special relativity. Hence we desire to bring their work to more widespread attention and to expand on it.

In Section III we construct eigenvectors of position and the position operator, following the work of Newton and Wigner [3]. The apparent difference between their position operator and the nonrelativistic case can be easily understood as being due to the fact that these two operators are constructed to act on different wavefunctions. They are, in fact, the same operator when compared using the same wavefunctions. This fact was also explained by Fong and Rowe [4].

For reference, we show how to construct positive energy solutions of the Klein-Gordon and Dirac equations in terms of probability amplitudes.

We consider the spinless case and find the Lorentz transformation properties of the position operator and derive some physically meaningful conclusions about the relativity of position measurements.

In Section IV we discuss the concept of causality in quantum mechanics, limited, as it is, by the uncertainty principle. In Section V we comment briefly on the relevance of our results to theories of interaction. Conclusions follow in Section VI.

Throughout this paper we use Heaviside-Lorentz units, in which $\hbar = c = \epsilon_0 = \mu_0 = 1$. We use the active convention for Poincaré transformations, in which, for example, the boost by velocity $\boldsymbol{\beta}_0$ of a particle of mass m_0 at rest produces a particle with momentum $p^\mu = m_0(\gamma_0, \gamma_0\boldsymbol{\beta}_0)^\mu$, with $\gamma_0 = 1/\sqrt{1 - \boldsymbol{\beta}_0^2}$.

II. RELATIVISTIC MOMENTUM-SPIN PROBABILITY AMPLITUDES FOR MASSIVE PARTICLES

To calculate the probability of a process in quantum mechanics, a number between 0 and 1, requires representing the physical states of the system by state vectors that have a finite normalization. The expressions for these probabilities take their simplest forms if the state vectors are normalized to unity, so we follow that convention throughout. We are interested in cases where one or more of the observables have eigenvalues with a continuous spectrum. Then the eigenvectors of such an observable are improper state vectors with delta function normalizations. To form a state vector with unit normalization requires constructing a coherent superposition of these basis vectors. From the complex number coefficients in this superposition can be identified the relativistic probability amplitudes. These amplitudes, the physical information they contain, the operators that act on them and the relativistic transformations of these quantities are the main topics dealt with here.

The concept of a fundamental particle is intimately linked to that of irreducible representations of the Poincaré group [3, 5, 6]. While we believe the electron (and positron) to have no substructure, to the best of our experimental knowledge, we take as their free momentum/spin-component basis vectors $|\,p, \frac{1}{2}, m\,\rangle$, labelled by the four-momentum, $p^\mu = (\omega, \boldsymbol{p})^\mu$, the spin, $1/2$, and the spin z component in the rest frame. These, as we will see, carry the irreducible representations of the Poincaré group. We leave off the charge and any other labels. (While we label the basis vectors by the four-momenta to make their transformations more transparent, we know that there are only three independent quantum numbers, the three components of $\boldsymbol{p}$.) If substructure were to be discovered, using an irreducible representation would become, at best, an approximation.

For a general spin, $s = 0, \frac{1}{2}, 1, \ldots$, we have basis vectors $|\,p, s, m\,\rangle$ with spin z component $m = -s, -(s-1), \ldots, s - 1, s$. These basis vectors are improper state vectors with delta function normalization. We choose to use the invariant normalization,

$$\langle\, p_1, s, m_1 \,|\, p_2, s, m_2 \,\rangle = \delta_{m_1 m_2} \omega_1 \delta^3(\boldsymbol{p}_1 - \boldsymbol{p}_2), \tag{2}$$

as this makes the following transformations take their simplest forms. These basis vectors carry irreducible unitary representations of spacetime translations [5, 6]:

$$U(T(a))\,|\,p, s, m\,\rangle = |\,p, s, m\,\rangle e^{+ip\cdot a}, \tag{3}$$

rotations (with double-valued representations for half-integral spins):

$$U(R)\,|\,p, s, m\,\rangle = \sum_{m'=-s}^{s} |\,Rp, s, m'\,\rangle \mathcal{D}^{(s)}_{m'm}(R) \tag{4}$$

and boosts:

$$U(\Lambda)\,|\,p, s, m\,\rangle = \sum_{m'=-s}^{s} |\,\Lambda p, s, m'\,\rangle \mathcal{W}^{(s)}_{m'm}(\Lambda p \leftarrow p), \tag{5}$$

the elements of the Poincaré group. In addition, they represent space inversion:

$$U(\mathcal{P})\,|\,(\omega, \boldsymbol{p}), s, m\,\rangle = \eta\,|\,(\omega, -\boldsymbol{p}), s, m\,\rangle \tag{6}$$

and time reversal

$$A(\mathcal{T})\,|\,(\omega, \boldsymbol{p}), s, m\,\rangle = |\,(\omega, -\boldsymbol{p}), s, -m\,\rangle(-)^{s-m}. \tag{7}$$

Here $\eta = \pm 1$ is called the intrinsic parity of the particle. Note that the antiunitary time reversal operator, $A(\mathcal{T})$, when applied to a superposition, involves taking the complex conjugate of the amplitudes. Also $\mathcal{W}$ is a matrix element of a Wigner rotation, which can be evaluated by

$$W(\Lambda p \leftarrow p) = \Lambda^{-1}[\Lambda p]\,\Lambda\,\Lambda[p]. \tag{8}$$

where

$$\Lambda[p] \equiv \Lambda(\frac{\boldsymbol{p}}{\omega}) \tag{9}$$

and $\Lambda(\boldsymbol{\beta})$ is a function of the boost velocity, $\boldsymbol{\beta}$. Explicit forms of the Wigner rotations are given in [6]. Two successive, noncollinear, boosts (from the rest momentum to p and then from p to Λp) produce a boost (from the rest momentum to Λp) preceded by a rotation in the rest frame. This is the physics behind the Thomas precession [7].

As discussed above, we must form coherent superpositions of the improper basis vectors to make normalized state vectors. We write the general case as

$$|\psi\rangle = \int d^3 p \sum_{m=-s}^{s} |p, s, m\rangle \frac{1}{\sqrt{\omega}} \Psi_m(p). \tag{10}$$

The factor of $1/\sqrt{\omega}$ is to compensate for the covariant normalization, making this expression just like the familiar nonrelativistic case. By this we mean that the state vectors

$$|p, s, m; \text{NR}\rangle = |p, s, m\rangle \frac{1}{\sqrt{\omega}} \tag{11}$$

have the familiar orthonormality

$$\langle p_1, s, m_1; \text{NR} | p_2, s, m_2; \text{NR}\rangle = \delta^3(\boldsymbol{p}_1 - \boldsymbol{p}_2) \tag{12}$$

used in nonrelativistic quantum mechanics. (Again we note that $\Psi_m(p)$ is a function of only three independent momentum components.)

We also define, for comparison, amplitudes

$$\Phi_m(p) = \sqrt{\omega}\, \Psi_m(p). \tag{13}$$

Now we can derive (not postulate) the relativistic transformation properties of the $\Psi_m(p)$ (and thus, of the $\Phi_m(p)$). The technique is to apply the unitary (or antiunitary) transformation to $|\psi\rangle$ and thus to the basis vectors, then manipulate the expression into the form

$$U/A|\psi\rangle = \int d^3 p \sum_{m=-s}^{s} |p, s, m\rangle \frac{1}{\sqrt{\omega}} \Psi'_m(p), \tag{14}$$

then extract the $\Psi'_m(p)$ by orthonormality.

The transformation rules for the Poincaré transformations are:

$$\text{Spacetime translations}: \quad \Psi'_m(p) = \Psi_m(p)\, e^{+ip\cdot a},$$

$$\text{Rotations}: \quad \Psi'_m(p) = \sum_{m'=-s}^{s} \mathcal{D}_{mm'}^{(s)}(R)\, \Psi_{m'}(R^{-1}p),$$

$$\text{Boosts}: \quad \Psi'_m(p) = \sqrt{\gamma_0(1 - \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})} \sum_{m'=-s}^{s} \mathcal{W}_{mm'}^{(s)}(p \leftarrow \Lambda^{-1}p)\Psi_{m'}(\Lambda^{-1}p), \tag{15}$$

where $\boldsymbol{\beta}_0$ is the boost velocity, $\gamma_0 = 1/\sqrt{1 - \boldsymbol{\beta}_0^2}$ and $\boldsymbol{\beta} = \boldsymbol{p}/\omega$ is the velocity of the particle. For the inversions, we have

$$\text{Space inversion}: \quad \Psi'_m(\omega, \boldsymbol{p}) = \eta\, \Psi_m(\omega, -\boldsymbol{p}),$$

$$\text{Time reversal}: \quad \Psi'_m(\omega, \boldsymbol{p}) = (-)^{s+m}\Psi^*_{-m}(\omega, -\boldsymbol{p}). \tag{16}$$

In these equations we have satisfied the requirements of special relativity. We have the rules for transforming our physical quantity under Poincaré transformations, rules which depend only on the translation, rotation and boost parameters. What we have here are examples of *nonmanifest* covariance.

These transformations all preserve the modulus squared of scalar products:

$$\left| \int d^3 p' \sum_{m'=-s}^{s} \Psi_{m'}^{(1)'*}(p')\Psi_{m'}^{(2)'}(p') \right|^2 = \left| \int d^3 p \sum_{m=-s}^{s} \Psi_m^{(1)*}(p)\Psi_m^{(2)}(p) \right|^2, \tag{17}$$

as can be easily verified from Eqs. (15,16). For example, we show the $s = 0$ boost case. With $p' = \Lambda p$, we have

$$\Psi^{(n)\prime}(p') = \sqrt{\gamma_0 (1 - \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})}\, \Psi^{(n)}(p) = \sqrt{\frac{\omega}{\omega'}}\, \Psi^{(n)}(p), \quad \text{for } n = 1, 2, \tag{18}$$

so

$$\int d^3 p'\, \Psi^{(1)\prime}(p')\Psi^{(2)\prime}(p') = \int \frac{d^3 p'}{\omega'}\, \omega\, \Psi^{(1)}(p)\Psi^{(2)}(p) = \int \frac{d^3 p}{\omega}\, \omega\, \Psi^{(1)}(p)\Psi^{(2)}(p) = \int d^3 p\, \Psi^{(1)}(p)\Psi^{(2)}(p). \tag{19}$$

We point out the important distinction between unitary covariance, manifest covariance and nonmanifest covariance for Poincaré transformations. Once the unitary Poincaré transformations of the basis vectors are specified (such as Eqs. (3,4,5,6,7)), the unitary transformations of all amplitudes and all observables written on those basis vectors are specified. Any equation involving observables, such as $i[A, B] = C$, will take the same form in all frames, $i[A', B'] = C'$, with $A' = U^\dagger(L)AU(L)$ etc. A subset of the observables will transform manifestly, as scalars, four-vectors or tensors. An example is the four-momentum operator, with $P'^\mu = U^\dagger(L)P^\mu U(L) = L^\mu{}_\nu P^\nu$. The remainder of the observables will transform nonmanifestly, not as scalars, four-vectors or tensors. The prime example of this will be the position operator (Eq. (67) below). Special relativity does not demand that all quantities of physical interest transform with Lorentz indices, merely that all such transformations are well defined and depend only on the translation, rotation and boost parameters.

The normalization condition becomes, in two forms,

$$1 = \int d^3 p \sum_{m=-s}^{s} |\Psi_m(p)|^2 = \int \frac{d^3 p}{\omega} \sum_{m=-s}^{s} |\Phi_m(p)|^2. \tag{20}$$

The expectation of the four-momentum operator, P^μ, becomes, also in two forms

$$\langle\, \psi\, |\, P^\mu\, |\, \psi\, \rangle = \int d^3 p\, p^\mu \sum_{m=-s}^{s} |\Psi_m(p)|^2 = \int \frac{d^3 p}{\omega}\, p^\mu \sum_{m=-s}^{s} |\Phi_m(p)|^2. \tag{21}$$

The first forms in both cases show the probability interpretation of our probability amplitudes, in exactly the same form as in the nonrelativistic theory. The quantity

$$\rho(\boldsymbol{p}) = \sum_{m=-s}^{s} |\Psi_m(p)|^2 \tag{22}$$

acts like a normalized momentum probability density, with $\rho(\boldsymbol{p})\Delta V(\boldsymbol{p})$ being the probability of measuring the momentum of the particle in a small volume $\Delta V(\boldsymbol{p})$ of momentum space. Furthermore, the use of a projector onto an eigenvector of momentum and spin component in

$$\langle\, \psi\, |\, p, s, m\, \rangle\langle\, p, s, m\, |\, \psi\, \rangle = |\Psi_m(p)|^2 \tag{23}$$

confirms the interpretation of $\Psi_m(p)$ as a position/spin-component probability amplitude.

The second forms in Eqs. (20,21) show manifestly the *invariance* of the normalization result and the *covariance* of the momentum expectation, once we recognize, as is easily seen from Eqs. (15) that the quantity

$$S(p) = \sum_{m=-s}^{s} |\Phi_m(p)|^2 \tag{24}$$

transforms as a scalar function under general Lorentz transformations, L:

$$S'(p) = S(L^{-1}p). \tag{25}$$

We are not saying that quantities like $\Phi_m(p)$ and the Dirac amplitudes (discussed below) might not be *useful* in constructing theories of interaction. We are pointing out that they are not probability amplitudes, while probability amplitudes $\Psi_m(p)$, with the same probability interpretation as in the nonrelativistic theory, can be formed and have well-defined relativistic transformation properties.

We will not investigate the relativity of spin measurements in this paper.

III.   THE POSITION OPERATOR AND ITS RELATIVISTIC TRANSFORMATIONS

A.   Position/spin-component probability amplitudes and the position operator

Newton and Wigner [3] proposed four conditions that should be satisfied by a state vector representing a massive particle, with spin, localized at the origin at time $t = 0$. The state vector is written as a superposition of the basis vectors $|\, p, s, m\,\rangle$. These conditions are

1. A spatial translation should produce a vector orthogonal to the original.

2. A linear superposition of localized state vectors produces another localized state vector.

3. A rotation should be represented by the linear combination of state vectors from a set: the set should carry a finite-dimensional irreducible representation of rotations. Also, the actions of space inversion and time reversal should also produce a state vector in this set.

4. The superposition coefficients should be everywhere regular functions of the momentum $\boldsymbol{p}$, which can be enforced by requiring finiteness upon the application of an infinitesimal boost, which involves a derivative with respect to momentum.

They found

$$|\, \boldsymbol{0}, s, m\,\rangle = \int \frac{d^3p}{\omega}|\, p, s, m\,\rangle\frac{\sqrt{\omega}}{(2\pi)^{\frac{3}{2}}}. \tag{26}$$

This spatially translates to

$$|\, \boldsymbol{x}, s, m\,\rangle = \int \frac{d^3p}{\sqrt{\omega}}|\, p, s, m\,\rangle\frac{e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}}{(2\pi)^{\frac{3}{2}}}, \tag{27}$$

and satisfies the first condition with the delta function normalization

$$\langle\, \boldsymbol{x}_1, s, m_1\,|\, \boldsymbol{x}_2, s, m_2\,\rangle = \delta_{m_1 m_2}\delta^3(\boldsymbol{x}_1 - \boldsymbol{x}_2). \tag{28}$$

Conditions 2 and 3 are clearly satisfied, with

$$U(R)|\, \boldsymbol{0}, s, m\,\rangle = \sum_{m'=-s}^{s}|\, \boldsymbol{0}, s, m'\,\rangle\mathcal{D}^{(s)}_{m'm}(R),$$

$$U(\mathcal{P})|\, \boldsymbol{0}, s, m\,\rangle = \eta\,|\, \boldsymbol{0}, s, m\,\rangle,$$

$$A(\mathcal{T})|\, \boldsymbol{0}, s, m\,\rangle = \sum_{m'=-s}^{s}|\, \boldsymbol{0}, s, -m\,\rangle(-)^{s-m}. \tag{29}$$

It can be easily verified that condition 4 is satisfied.

Note that we use the same label, m, for the spin component in each rest frame and the spin component at position $\boldsymbol{x}$. Clearly spin measurements commute with position measurements and momentum measurements.

From Eq. (27) we see that we have exactly the same position eigenvectors as in the nonrelativistic theory, since again the factor of $1/\sqrt{\omega}$ compensates for the covariant normalization. Then we can define amplitudes that we expect to have an interpretation as position/spin-component probability amplitudes

$$\psi_m(\boldsymbol{x}) = \langle\, \boldsymbol{x}, s, m\,|\, \psi\,\rangle = \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}}\Psi_m(p)\, e^{+i\boldsymbol{p}\cdot\boldsymbol{x}}. \tag{30}$$

In the Schrödinger picture we have time-dependent amplitudes

$$\psi_m(t, \boldsymbol{x}) = \langle\, \boldsymbol{x}, s, m\,|\, e^{-iHt}\,|\, \psi\,\rangle = \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}}\Psi_m(p)\, e^{+i(\boldsymbol{p}\cdot\boldsymbol{x}-\omega t)}. \tag{31}$$

The time evolution of this wavefunction is governed by the (free) relativistic Schrödinger equation

$$i\frac{\partial}{\partial t}\psi_m(t, \boldsymbol{x}) = \hat{H}\,\psi_m(t, \boldsymbol{x}). \tag{32}$$

The Hamiltonian is completely defined in momentum space, where it has only positive eigenvalues,

$$\hat{H}\,|\,p,s,m\,\rangle = \omega\,|\,p,s,m\,\rangle = +\sqrt{\boldsymbol{p}^2 + m_0^2}\,|\,p,s,m\,\rangle. \tag{33}$$

It does not have a representation as a differential operator in position space except in the nonrelativistic limit, where it reduces to

$$\hat{H} \to m_0 - \frac{1}{2m_0}\nabla^2. \tag{34}$$

It is perhaps easiest to understand in position space in terms of Fourier transforms:

$$\hat{H} = \hat{\mathcal{F}}^*(x \leftarrow p)\,\sqrt{\boldsymbol{p}^2 + m_0^2}\,\hat{\mathcal{F}}(p \leftarrow x), \tag{35}$$

where

$$\hat{\mathcal{F}}(p \leftarrow x)\,f(x) = \int \frac{d^3x}{(2\pi)^{\frac{3}{2}}} e^{ip\cdot x}\,f(x),$$

$$\hat{\mathcal{F}}^*(x \leftarrow p)\,F(p) = \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}} e^{-ip\cdot x}\,F(p). \tag{36}$$

We confirm the covariance of the relativistic Schrödinger equation under boosts. With

$$\psi'_{m'}(x') = \langle\,\boldsymbol{x}',m'\,|\,e^{-iHt'}\,U(\Lambda)\,|\,\psi\,\rangle = \int \frac{d^3p'}{(2\pi)^{\frac{3}{2}}}\,\Psi'_{m'}(p')\,e^{-ip'\cdot x'}, \tag{37}$$

we have

$$i\frac{\partial}{\partial t'}\psi'_{m'}(x') = \int \frac{d^3p'}{(2\pi)^{\frac{3}{2}}}\,\omega'\,\Psi'_{m'}(p')\,e^{-ip'\cdot x'}$$

$$= \hat{\mathcal{F}}^*(x' \leftarrow p')\,\omega'\,\hat{\mathcal{F}}(p' \leftarrow x')\,\psi'_{m'}(x'). \tag{38}$$

So

$$i\frac{\partial}{\partial t'}\psi'_{m'}(t',\boldsymbol{x}') = \hat{H}'\,\psi'_{m'}(t',\boldsymbol{x}'), \tag{39}$$

with

$$\hat{H}' = \hat{\mathcal{F}}^*(x' \leftarrow p')\,\sqrt{\boldsymbol{p}'^2 + m_0^2}\,\hat{\mathcal{F}}(p' \leftarrow x'). \tag{40}$$

The relativistic Schrödinger equation, Eq. (32), can be seen as the zero component of the equation

$$i\partial^\mu\psi_m(x) = \hat{P}^\mu\psi_m(x), \tag{41}$$

with

$$\hat{\boldsymbol{P}} = \hat{\mathcal{F}}^*(x \leftarrow p)\,\boldsymbol{p}\,\hat{\mathcal{F}}(p \leftarrow x). \tag{42}$$

We note that the position probability amplitude also satisfies the Klein-Gordon equation (1).

Since Eq. (31) is a Fourier transform, the Parseval theorem tells us that the normalization condition can be written

$$\int d^3x \sum_{m=-s}^{s} |\psi_m(t,x)|^2 = \int d^3p \sum_{m=-s}^{s} |\Psi_m(p)|^2 = 1, \tag{43}$$

which is a Poincaré (and inversion) invariant statement, as we have already seen that fact for the momentum integral.

Then we can find the Poincaré transformation (and inversion) properties of these amplitudes:

$$\text{Spacetime Translations}: \quad \psi'_m(x) = \psi_m(x - a),$$

$$\text{Rotations}: \quad \psi'_m(x) = \sum_{m'=-s}^{s} \mathcal{D}^{(s)}_{mm'}(R)\psi_{m'}(R^{-1}x),$$

$$\text{Boosts}: \quad \psi'_m(x) = \sum_{m'=-s}^{s} \hat{B}_{mm'}\psi_{m'}(\Lambda^{-1}x),$$

$$\text{Space inversion}: \quad \psi'_m(t, \boldsymbol{x}) = \eta\, \psi_m(t, -\boldsymbol{x}),$$

$$\text{Time reversal}: \quad \psi'_m(t, \boldsymbol{x}) = (-)^{s+m}\, \psi^*_{-m}(-t, \boldsymbol{x}). \tag{44}$$

The operator appearing in the boost transformation is, in terms of Fourier transforms,

$$\hat{B}_{mm'} = \hat{\mathcal{F}}^*(x \leftarrow p)\, \sqrt{\gamma_0(1 - \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})}\, \mathcal{W}^{(s)}_{mm'}(p \leftarrow \Lambda^{-1}p)\, \hat{\mathcal{F}}(\Lambda^{-1}p \leftarrow \Lambda^{-1}x). \tag{45}$$

This transformation is nonlocal, as it must be to preserve probability.

These transformations all preserve the modulus squared of scalar products:

$$\left| \int d^3x' \sum_{m'=-s}^{s} \psi^{(1)'*}_m(x')\psi^{(2)'}_m(x') \right|^2 = \left| \int d^3x \sum_{m=-s}^{s} \psi^{(1)*}_m(x)\psi^{(2)}_m(x) \right|^2,$$

as can be easily verified from Eqs. (44).

Since at $t = 0$ the Fourier transform inverse to Eq. (30) is

$$\Psi_m(p) = \int \frac{d^2x}{(2\pi)^{\frac{3}{2}}} \psi_m(0, \boldsymbol{x})\, e^{-i\boldsymbol{p}\cdot\boldsymbol{x}}, \tag{46}$$

the operator in momentum space that produces a factor of $\boldsymbol{x}$, the position operator at $t = 0$, is simply

$$\hat{\boldsymbol{x}} = i\frac{\partial}{\partial\boldsymbol{p}}, \tag{47}$$

as in the nonrelativistic theory, where $\partial/\partial\boldsymbol{p} \equiv (\partial/\partial p_x, \partial/\partial p_y, \partial/\partial p_z)$. Since $i[\hat{H}, \hat{\boldsymbol{x}}] = \hat{\boldsymbol{\beta}}$ and $i[\hat{H}, \hat{\boldsymbol{\beta}}] = 0$, we have, in the Heisenberg picture,

$$\hat{\boldsymbol{x}}(t) = e^{+iHt}\, \hat{\boldsymbol{x}}\, e^{-iHt} = \hat{\boldsymbol{x}} + \hat{\boldsymbol{\beta}}t, \tag{48}$$

where $\hat{\boldsymbol{\beta}}$ is the operator with eigenvalue $\boldsymbol{p}/\omega$ acting on a momentum/spin-component eigenvector.

Newton and Wigner [3] find the position operator to be

$$\hat{\boldsymbol{x}}_{\text{NW}} = i\frac{\partial}{\partial\boldsymbol{p}} - i\frac{\boldsymbol{p}}{2\omega^2}, \tag{49}$$

but theirs is constructed to act on $\Phi_m(p)$ wavefunctions rather than $\Psi_m(p)$ probability amplitudes, with the invariant measure in their scalar products. The apparent difference is resolved in the identity

$$\int \frac{d^3p}{\omega} \sum_{m=-s}^{s} \Phi^{(1)*}_m(p)\{i\frac{\partial}{\partial\boldsymbol{p}} - i\frac{\boldsymbol{p}}{2\omega^2}\}\Phi^{(2)}_m(p) = \int d^3p \sum_{m=-s}^{s} \Psi^{(1)*}_m(p)\{i\frac{\partial}{\partial\boldsymbol{p}}\}\Psi^{(2)}_m(p). \tag{50}$$

The Hermiticity of the position operator is most easily confirmed in the second form.

The use of a projector onto an eigenvector of position and spin component in (Schrödinger picture)

$$\langle\, \psi(t)\,|\,\boldsymbol{x}, m\,\rangle\langle\,\boldsymbol{x}, m\,|\,\psi(t)\,\rangle = |\psi_m(t, x)|^2 \tag{51}$$

confirms the interpretation of $\psi_m(x)$ as a position/spin-component probability amplitude.

Note that we deal only with positive energy solutions of the relativistic Schrödinger equation, Eq. (32), and that the action of the position operator does not introduce negative energies. This was discussed by Newton and Wigner [3], who pointed out that the argument that measuring the position of an electron in a very small volume would lead

to pair creation is not consistent with this construction of the position operator. While pair creation would occur if high energy photons were used to localize the electron, the simplest way to localize an electron in a very small volume is to observe it from a frame boosted by a highly relativistic velocity relative to the average rest frame. Lorentz contraction will localize the electron in an arbitrarily small volume without pair creation, as can be verified in the rest frame. Lorentz contraction and the spreading with time of relativistic wavepackets will be dealt with in a paper in preparation.

We also note that our results for spin other than zero are different from those of Newton and Wigner [3], since they derived the action of their position operator on Dirac amplitudes.

B. Locally transforming position amplitudes

From the probability amplitude, $\Psi(p)$, for a massive, spinless particle, we can construct an amplitude that transforms locally as a scalar function of $x^\mu = (t, \boldsymbol{x})^\mu$,

$$\varphi(x) = \int \frac{d^3p}{\omega} e^{-ip\cdot x} \sqrt{\omega}\, \Psi(p),$$

and is a positive energy solution of the Klein-Gordon equation.

From the probability amplitudes, $\Psi_{\pm\frac{1}{2}}(p)$, for an electron, we can construct a positive energy solution of the Dirac equation,

$$\sum_{b=1}^{4} [i\gamma^\mu \partial_\mu - m_e]_{ab} \Psi_b^{\text{Dirac}}(x) = 0 \quad \text{for } a = 1, \ldots, 4,$$

with the gamma matrices (γ^μ) in the Weyl or chiral representation [2]. This is the four-component column vector

$$\Psi^{\text{Dirac}}(x) = \int \frac{d^3p}{\omega} e^{-ip\cdot x} \frac{1}{\sqrt{2}} \begin{pmatrix} D^{(-)}_{+\frac{1}{2},m}[p] \\ D^{(-)}_{-\frac{1}{2},m}[p] \\ D^{(+)}_{+\frac{1}{2},m}[p] \\ D^{(+)}_{-\frac{1}{2},m}[p] \end{pmatrix} \sqrt{\omega}\, \Psi_m(p),$$

where

$$D^{(r)}[p] = \sqrt{\frac{\omega + m_e}{2m_e}} + r\sqrt{\frac{\omega - m_e}{2m_e}} \hat{\boldsymbol{p}} \cdot \boldsymbol{\sigma} \quad \text{for } r = \pm 1$$

and $\boldsymbol{\sigma}$ are the 2×2 Pauli matrices. These components transform locally under a four-dimensional nonunitary transformation. Weinberg [8] shows how to construct locally transforming amplitudes for any spin, but his interest was not in probability amplitudes.

C. The boost transformation of the position operator

Note that the momentum dependence of the Wigner rotation (the boost case of Eq. (15)) means that the position wavefunction of a particle with spin will change under boosts in a way different from the spinless case. We restrict our attention here to the spinless case and derive the boost transformation of the position operator.

In what follows we neglect to put hats on operators, using them instead for unit vectors. The unitary boost transformation can be written

$$U(\Lambda) = e^{-i\boldsymbol{\zeta}\cdot\boldsymbol{K}}, \tag{52}$$

where $\boldsymbol{\zeta}$ is called the rapidity, satisfying

$$\cosh\zeta = \gamma_0, \quad \sinh\zeta\, \hat{\boldsymbol{\zeta}} = \gamma_0\boldsymbol{\beta}_0, \tag{53}$$

and the boost generator is

$$\boldsymbol{K} = -\frac{i}{2}\omega\frac{\partial}{\partial\boldsymbol{p}} - \frac{i}{2}\frac{\partial}{\partial\boldsymbol{p}}\omega = -\frac{1}{2}\{\omega, \boldsymbol{x}\}, \tag{54}$$

found by considering Eq. (15) for an infinitesimal transformation. Here $\boldsymbol{K}$ is the operator defined to act on the $\Psi(p)$ wavefunctions. (The anticommutator of two Hermitian operators is defined as $\{A, B\} = AB + BA$ and is clearly Hermitian.)

We start by separating the position operator into parts parallel and perpendicular to the boost direction:

$$\boldsymbol{x} = \boldsymbol{x}_\| + \boldsymbol{x}_\perp, \tag{55}$$

with

$$\boldsymbol{x}_\| = \boldsymbol{x} \cdot \hat{\boldsymbol{\zeta}} \hat{\boldsymbol{\zeta}}. \tag{56}$$

We note that

$$\hat{\boldsymbol{\zeta}} \boldsymbol{\zeta} \cdot \boldsymbol{K} = -\frac{\zeta}{2}\{\omega, \boldsymbol{x}_\|\}, \tag{57}$$

and this operator will clearly be invariant under this Lorentz transformation. This gives

$$\{\omega', \boldsymbol{x}_\|'\} = \{\omega, \boldsymbol{x}_\|\}. \tag{58}$$

This equation is easily solved to give

$$\boldsymbol{x}_\|' = \frac{1}{2}\left\{\frac{1}{\gamma_0(1 + \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})}, \boldsymbol{x}_\|\right\}, \tag{59}$$

in manifestly Hermitian form.

For the perpendicular components, we define two unit vectors so that $\{\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2, \hat{\boldsymbol{\zeta}}\}$ is a right-handed set of three mutually perpendicular axes. Then we calculate

$$\hat{\boldsymbol{u}}_1 \cdot \boldsymbol{K}' = e^{+i\boldsymbol{\zeta}\cdot\boldsymbol{K}}\, \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{K}\, e^{-i\boldsymbol{\zeta}\cdot\boldsymbol{K}} = \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{K} + i\zeta[\hat{\boldsymbol{\zeta}} \cdot \boldsymbol{K}, \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{K}] - \frac{1}{2}\zeta^2[\hat{\boldsymbol{\zeta}} \cdot \boldsymbol{K}, [\hat{\boldsymbol{\zeta}} \cdot \boldsymbol{K}, \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{K}]] + \dots \tag{60}$$

by using the Lorentz group commutators [7]

$$[K_i, K_j] = -i\epsilon_{ijk}J_k, \quad [K_i, J_j] = +i\epsilon_{ijk}K_k. \tag{61}$$

A similar calculation gives $\hat{\boldsymbol{u}}_2 \cdot \boldsymbol{K}'$. We find

$$\hat{\boldsymbol{u}}_1 \cdot \boldsymbol{K}' = \gamma_0\, \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{K} + \gamma_0\beta_0\, \hat{\boldsymbol{u}}_2 \cdot \boldsymbol{J},$$
$$\hat{\boldsymbol{u}}_2 \cdot \boldsymbol{K}' = \gamma_0\, \hat{\boldsymbol{u}}_2 \cdot \boldsymbol{K} - \gamma_0\beta_0\, \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{J}. \tag{62}$$

Using $\boldsymbol{J} = \boldsymbol{x} \times \boldsymbol{p}$, we find

$$\hat{\boldsymbol{u}}_2 \cdot \boldsymbol{J} = \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{p}\, \hat{\boldsymbol{\zeta}} \cdot \boldsymbol{x} - p_\|\, \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{x},$$
$$-\hat{\boldsymbol{u}}_1 \cdot \boldsymbol{J} = \hat{\boldsymbol{u}}_2 \cdot \boldsymbol{p}\, \hat{\boldsymbol{\zeta}} \cdot \boldsymbol{x} - p_\|\, \hat{\boldsymbol{u}}_2 \cdot \boldsymbol{x}. \tag{63}$$

Together, these results give

$$\boldsymbol{K}_\perp' = \gamma_0\boldsymbol{K}_\perp + \gamma_0(\boldsymbol{p}_\perp\,\boldsymbol{\beta}_0 \cdot \boldsymbol{x} - \boldsymbol{\beta}_0 \cdot \boldsymbol{p}\,\boldsymbol{x}_\perp). \tag{64}$$

Solving this equation for $\boldsymbol{x}_\perp'$, using

$$\boldsymbol{K}_\perp' = -\frac{1}{2}\{\omega, \boldsymbol{x}_\perp'\}, \tag{65}$$

etc. gives

$$\boldsymbol{x}_\perp' = \boldsymbol{x}_\perp - \frac{1}{2}\left\{\frac{1}{1 + \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta}}, \boldsymbol{\beta}_\perp\,\boldsymbol{\beta}_0 \cdot \boldsymbol{x}\right\}. \tag{66}$$

So the transformation law is

$$\boldsymbol{x}' = \boldsymbol{x}_\perp - \frac{1}{2}\left\{\frac{1}{1 + \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta}}, \boldsymbol{\beta}_\perp\,\boldsymbol{\beta}_0 \cdot \boldsymbol{x}\right\} + \frac{1}{2}\left\{\frac{1}{\gamma_0(1 + \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})}, \boldsymbol{x}_\|\right\}. \tag{67}$$

We know that the definition of velocity, $\boldsymbol{\beta} = i[H, \boldsymbol{x}]$, must take the same form in all frames. However, it was of value to verify this result using the explicit form of the transformation of the position operator, as a check on Eq. (67). We found

$$i[H', \boldsymbol{x}'] = i[\gamma_0(H + \boldsymbol{\beta}_0 \cdot \boldsymbol{P}), \boldsymbol{x}'] = \frac{\boldsymbol{\beta}_\perp + \gamma_0(\boldsymbol{\beta}_\| + \boldsymbol{\beta}_0)}{\gamma_0(1 + \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})} = \boldsymbol{\beta}', \tag{68}$$

the law for the relativistic transformation of velocity [9].

Then we verified the *invariance* of the fundamental commutator between position and momentum

$$[x'_i, P'_j] = [x_i, P_j] = i\delta_{ij}. \tag{69}$$

Because the position and momentum have indices, we might have thought that this commutator transformed like components of a tensor.

D. Lorentz transformations of "average events"

We examine the concept of an event in this theory. Since there is no time operator in quantum mechanics, the position operator cannot transform as the spatial components of a four-vector. We have derived its transformation formula and it is linear and homogeneous in the position operator components with nothing like a time operator appearing. What, though, of expectation values? If we define

$$x^\mu = (t, \langle \psi \,|\, \boldsymbol{x}(t) \,|\, \psi \rangle)^\mu \quad \text{and} \quad x'^\mu = (t', \langle \psi \,|\, \boldsymbol{x}'(t') \,|\, \psi \rangle)^\mu \tag{70}$$

with

$$t' = \gamma_0(t + \boldsymbol{\beta}_0 \cdot \langle \psi \,|\, \boldsymbol{x}(t) \,|\, \psi \rangle) \tag{71}$$

and

$$\boldsymbol{x}'(0) = U^\dagger(\Lambda)\, \boldsymbol{x}(0)\, U(\Lambda), \tag{72}$$

do we find

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu? \tag{73}$$

Here, in the Heisenberg picture, we have

$$\boldsymbol{x}(t) = e^{+iHt}\, \boldsymbol{x}\, e^{-iHt} = \boldsymbol{x} + \boldsymbol{\beta}t \quad \text{and} \quad \boldsymbol{x}'(t') = e^{+iHt'}\, \boldsymbol{x}'\, e^{-iHt'} = \boldsymbol{x}' + \boldsymbol{\beta}'t', \tag{74}$$

with $\boldsymbol{\beta}'$ as in Eq. (68).

We found the result to be correct, but only as a good approximation for wavepackets that are narrow in momentum space. The proof involves replacing operators by their expectation values, which is seen to be valid to a good approximation using wavepackets of this form. A typical term to be considered is

$$\boldsymbol{T} = \langle \psi \,|\, \frac{1}{2} \left\{ \frac{1}{\gamma_0(1 + \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})}, \boldsymbol{x}_\| \right\} \,|\, \psi \rangle. \tag{75}$$

After an integration by parts in which the boundary terms are required to vanish, this takes the form

$$\boldsymbol{T} = \int d^3p\, \frac{1}{2} \left\{ \Psi^*(\boldsymbol{p})\, f(\boldsymbol{p}) \left\{ i\frac{\partial}{\partial \boldsymbol{p}_\|}\Psi(\boldsymbol{p}) \right\} - \left\{ i\frac{\partial}{\partial \boldsymbol{p}_\|}\Psi^*(\boldsymbol{p}) \right\} f(\boldsymbol{p})\, \Psi(\boldsymbol{p}) \right\}, \tag{76}$$

with

$$f(\boldsymbol{p}) = \frac{1}{\gamma_0(1 + \boldsymbol{\beta}_0 \cdot \boldsymbol{\beta})}. \tag{77}$$

To proceed, we consider Gaussian wavepackets of the form

$$\Psi(\boldsymbol{p}) = \frac{e^{-|\boldsymbol{p}-\bar{\boldsymbol{p}}|^2/4\sigma_p^2}}{(2\pi\sigma_p^2)^{\frac{3}{4}}} e^{-i\boldsymbol{p}\cdot\bar{\boldsymbol{x}}}. \tag{78}$$

A more general result would be desirable, but is not available at this point. Since

$$i\frac{\partial}{\partial \boldsymbol{p}}\Psi(p) = \{-i\frac{\boldsymbol{p}-\bar{\boldsymbol{p}}}{2\sigma_p^2} + \bar{\boldsymbol{x}}\}\Psi(p),$$

(79)

this leads to

$$\boldsymbol{T} = \bar{\boldsymbol{x}}_\parallel \int d^3p\, |\Psi(p)|^2 f(\boldsymbol{p}).$$

(80)

Since the momentum probability density is a narrow function of momentum peaked at $\boldsymbol{p} = \bar{\boldsymbol{p}}$, while $f(\boldsymbol{p})$ is slowly varying on $|\boldsymbol{p}-\bar{\boldsymbol{p}}| \lesssim \sigma_p$, we expand the latter in powers of $\boldsymbol{p} - \bar{\boldsymbol{p}}$. The integral of the first-order term vanishes by the spherical symmetry of the momentum probability density around $\boldsymbol{p} = \bar{\boldsymbol{p}}$. Then we find an upper bound on the fractional remainder, ε, in

$$\boldsymbol{T} = \frac{\bar{\boldsymbol{x}}_\parallel}{\gamma_0(1 + \boldsymbol{\beta}_0 \cdot \bar{\boldsymbol{\beta}})}\{1 + \varepsilon\}$$

(81)

(with $\bar{\boldsymbol{\beta}} = \bar{\boldsymbol{p}}/\sqrt{\bar{\boldsymbol{p}}^2 + m_0^2}$) of

$$|\varepsilon| \lesssim \bar{\boldsymbol{\beta}}^2 \left(\frac{\sigma_p}{|\bar{\boldsymbol{p}}|}\right)^2,$$

(82)

required to be much less than unity.

It is important to note that these results on two-component relativistic probability amplitudes for electrons and positrons in no way contradict Dirac's construction [2] of a Hermitian current *operator* for electrons and positrons, which transforms locally as a four-vector function of position and time and has a zero component that is negative definite (electrons) or positive definite (positrons). We note that it was necessary in the construction to use locally transforming *four*-component objects. These objects are not probability amplitudes.

Similarly, the four-component Dirac *amplitudes* are very useful in solving the hydrogen atom spectrum [10]. However, the physical meaning of the components is unclear, their values change in different gamma matrix representations and they are clearly not probability amplitudes.

In fact these results complement each other. To find the expectation value of the Dirac current (which is written on the positive energy $|p, \frac{1}{2}, m\rangle$ basis) and to be able to interpret the results requires use of the relativistic probability amplitudes to form normalized state vectors. There is no need to decide whether position measurements or charge density measurements for an electron are the "right" measurements relativistically. They must be seen as complementary measurements, with the construction of the detectors expected to be quite different in the two cases.

IV. CAUSALITY IN QUANTUM MECHANICS

The concept of causality in relativistic quantum mechanics is limited by the uncertainty principle and the inherent probabilistic nature of the quantum world. A particle localized at a point at a particular time is represented by an improper state vector, not part of the physical Hilbert space. A particle in a physical state always has a nonvanishing spatial extent over which it can be detected. Thus it is not possible to define a light cone with exact boundaries for physical particle events.

The probability distribution in the velocity, $\boldsymbol{\beta}$, of a massive particle vanishes for $|\boldsymbol{\beta}| \geq 1$. So individual measurements of speed will always give a result less than the speed of light, and the expectation value of velocity will always be less than unity in magnitude. These results are reassuring, but stronger definitions of causality have been proposed and found to be violated.

Hegerfeldt's criteria [11, 12] for causality ask that the *probability density* propagate at less than the speed of light. The position probability amplitudes presented here fail his test. We acknowledge that some readers will consider this violation to be a flaw in our theory. Example position probability densities have been presented [13] that violate this concept of causality quite severely.

These violating state vectors are typically such that the expectations of higher powers of $|\boldsymbol{x}|$ or $|\boldsymbol{p}|$ are infinite, and as such may not be physically realizable. More research needs to be done on the shapes of wavepackets that can be produced in nature, for example from radioactive decay. Experimental verification of a small violation of a classical causality condition would be of great interest. We argue that this would be received as a feature of probabilistic quantum mechanics rather than a failure of relativity.
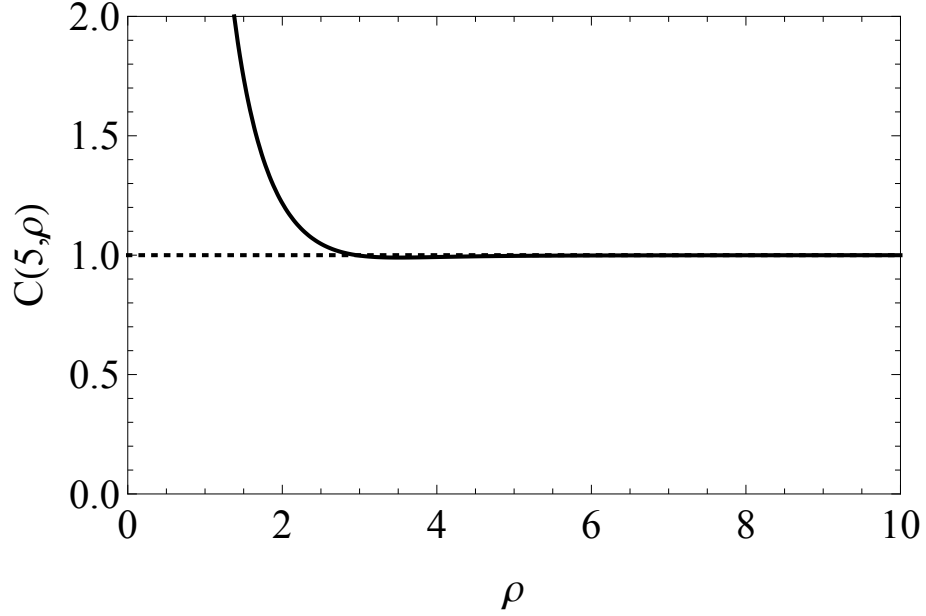
Figure 1. The causality ratio $C(5, \rho)$.

We argue that the Newton-Wigner definition of the position operator is the only correct choice and that Hegerfeldt's criteria for causality are too strict. In a probabilistic theory, it should be sufficient that a classical concept of causality be only violated with small probabilities.

An example illustrates this point. We consider a Gaussian momentum wavefunction

$$\Psi(\boldsymbol{\kappa}) = \frac{e^{-\boldsymbol{\kappa}^2/4}}{(2\pi)^{\frac{3}{4}}},$$

normalized to $\int d^3\kappa \, |\Psi(\boldsymbol{\kappa})|^2 = 1$, with the dimensionless scaled momentum $\boldsymbol{\kappa} = \boldsymbol{p}/\sigma_p$. We choose the momentum width $\sigma_p \gg m_0$, giving a very small spatial width ($\sigma_x = 1/2\sigma_p \ll 1/m_0$, much smaller than the Compton wavelength) at $t = 0$. This approximates the massless case and the superposition is dominated by highly relativistic momenta. We find the approximate time-dependent spatial wavefunction, spherically symmetric [14] (3.462.1, 9.246 and 9.248.1),

$$\psi(\tau, \boldsymbol{\rho}) = \frac{-i}{(2\pi)^{\frac{3}{4}}} \frac{1}{\sqrt{\pi}} \frac{1}{\rho} \{e^{-(\tau-\rho)^2/8} D_{-2}(i(\tau-\rho)/\sqrt{2}) - e^{-(\tau+\rho)^2/8} D_{-2}(i(\tau+\rho)/\sqrt{2})\}.$$

normalized to $\int d^3\rho \, |\psi(\tau, \boldsymbol{\rho})|^2 = 1$, with $\boldsymbol{\rho} = \boldsymbol{r}/\sigma_x$ and $\tau = t/\sigma_x$. The asymptotic form of this wavefunction as $\rho \to \infty$ is

$$\psi(\tau, \boldsymbol{\rho}) \to \frac{1}{(2\pi)^{\frac{3}{4}}} e^{-\rho^2/4},$$

which is stronger localization than was considered in Hegerfeldt's second theorem [12]. The expectations of all powers of $|\boldsymbol{\rho}|$ and $|\boldsymbol{\kappa}|$ are finite for these wavefunctions.

We invoke a causality test very similar to that proposed by Rosenstein and Usher [13]. We define the causality ratio, for $\tau > 0$,

$$C(\tau, \rho) = \frac{\int_0^{\rho+\tau} 4\pi\rho'^2 d\rho' \, |\psi(\tau, \boldsymbol{\rho}')|^2}{\int_0^\rho 4\pi\rho'^2 d\rho' \, |\psi(0, \boldsymbol{\rho}')|^2}$$

and require it to be everywhere greater than or equal to unity to satisfy classical causality, so that no probability leaks out of the light cone. Numerical results are shown in Figure 1.

While we would say that the criterion appears to be well satisfied, it is relevant to note that there is actually a region of tiny violations. We find $C(5, 5) = 0.996958$. So even this particularly well-behaved wavepacket fails the strict application of a causality criterion, but would certainly pass a looser, probabilistic interpretation that we argue is required of quantum mechanics.

V. INTERACTION

Of course everything presented above is for *free* particles, or particles isolated far from other particles with which they might interact. It is not our intention in this paper to discuss how to construct a relativistic theory of interactions involving these probability amplitudes. Many would say that we already have a relativistic theory of interaction in relativistic quantum field theory (QFT). We argue, though, that probability amplitudes are a necessary *addition* to quantum field theory. That theory calculates elements of the S matrix between plane wave states (momentum eigenvectors), of the example form

$$\mathcal{M} = \langle\, p_3, m_3; p_4, m_4 \,|\, S \,|\, p_1, m_1; p_2, m_2 \,\rangle. \tag{83}$$

To calculate a probability, between 0 and 1, requires construction of normalized state vectors, which must involve relativistic probability amplitudes. Thus an example probability would be

$$P = \int \frac{d^3 p_1}{\sqrt{\omega_1}} \, \Psi_{m_1}(p_1) \int \frac{d^3 p_2}{\sqrt{\omega_2}} \, \Psi_{m_2}(p_2) \int \frac{d^3 p_3}{\sqrt{\omega_3}} \, \Psi^*_{m_3}(p_3) \int \frac{d^3 p_4}{\sqrt{\omega_4}} \, \Psi^*_{m_4}(p_4) \langle\, p_3, m_3; p_4, m_4 \,|\, S \,|\, p_1, m_1; p_2, m_2 \,\rangle. \tag{84}$$

Wavepackets can regulate divergences. This author presented a wavepacket treatment of Coulomb scattering [15]. The use of wavepackets introduces a convergence factor into the partial wave series, which otherwise diverges for all scattering angles.

We are not saying that wavepackets alone can tame the infinities of relativistic quantum field theory. In Eq. (84) above, we assume that the S matrix element has already been renormalized before applying the wavepacket superposition. The infinities of QFT are more deep-rooted than that.

We do present an argument for the consideration of the reader. Suppose, in a purely quantum mechanical treatment, that we have a Hamiltonian that can be diagonalized, with finite energies. Then the evolution operator, an exponential of the Hamiltonian, is also diagonal in this basis, and is unitary with finite matrix elements. Then the central problem becomes how to construct incoming and outgoing normalized state vectors to represent particles localized far from each other. These state vectors would be written as superpositions of the eigenvectors of the Hamiltonian, using relativistic probability amplitudes. If this can be done, we can evolve the incoming state vector in time, then calculate the probability

$$P = |\langle\, \text{out} \,|\, e^{-iHT} \,|\, \text{in} \,\rangle|^2.$$

From unitarity, the normalization of the state vectors and Schwartz's inequality, this probability is guaranteed to be finite, and between 0 and 1.

In the opinion of this author, we do not yet have a quantum mechanical, relativistic and *finite* theory of interaction with these properties.

VI. CONCLUSIONS

We have defined momentum/spin-component probability amplitudes and position/spin-component probability amplitudes for a massive particle of general spin. We have found their transformation properties under spacetime translations, general Lorentz transformations, space inversion and time reversal. We have defined the position operator and derived its relativistic transformation properties. The results are all very close to what is done in nonrelativistic quantum mechanics. This should come as no surprise, since any relativistic theory must reduce to the nonrelativistic form for small velocities.

We discussed the limitations on the concept of causality imposed by the uncertainty principle.

We discussed how relativistic probability amptides must be a part of any theory of interaction, including quantum field theory.

---

[1] L. P. Horwitz, *Relativistic Quantum Mechanics* (Springer, Dordrecht, 2015).
[2] C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, 1st ed. (McGraw-Hill Inc., 1980).
[3] T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).
[4] R. Fong and E. Rowe, Annals of Physics **46**, 559 (1968).
[5] E. Wigner, Annals of Mathematics **40**, 149 (1939).

[6] F. R. Halpern, *Special Relativity and Quantum Mechanics*, 1st ed. (Prentice-Hall, Inc., N.J., 1968).

[7] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (John Wiley and Sons, Inc., 1975).

[8] S. Weinberg, Phys. Rev. **133**, B1318 (1964).

[9] W. Rindler, *Essential Relativity: Special, General, and Cosmological*, 2nd ed. (Springer-Verlag N.Y., 1977).

[10] A. Messiah, *Quantum Mechanics*, Vol. 1 and 2 (North-Holland, Amsterdam and John Wiley and Sons, N.Y., 1961).

[11] G. C. Hegerfeldt, Phys. Rev. D **10**, 3320 (1974).

[12] G. C. Hegerfeldt, Phys. Rev. Lett. **54**, 2395 (1985).

[13] B. Rosenstein and M. Usher, Phys. Rev. D **36**, 2381 (1987).

[14] I. S. Gradsteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products*, corrected and enlarged ed. (Academic Press, Inc., San Diego, CA, 1980).

[15] S. E. Hoffmann, J. Phys. B: At. Mol. Opt. Phys. **50**, 215302 (2017).