On Newton–Wigner Localization 
and Superluminal Propagation Speeds

S. N. M. Ruijenaars*

Department of Mathematics, Texas A & M University. 
College Station, Texas 77843

Received May 4, 1981; revised July 1, 1981

We study the acausal behavior of free massive relativistic particles, as predicted by the Newton–Wigner position operator. The main result is an upper bound for the probability of superluminal propagation speeds. This bound implies the detection probability of acausal events is vanishingly small under present laboratory conditions.

1. Introduction

It is surprising that more than 50 years after the invention of quantum mechanics a satisfactory understanding of the quantum measurement process is still lacking. Progress has been made in the description of the measurement of observables with discrete spectra (for a recent review, see, e.g., Ref. [1]), but a comparable understanding has not been reached for an observable like position. This is so in spite of the fact that the approximate determination of the position of a physical system may be regarded as the most fundamental measurement of all. Indeed, a terrestrial measurement of any other observable will not meet with much success if the system resides on the Moon.

Difficulties with quantum localization occur already at the nonrelativistic level. We will begin by discussing this in more detail. Firstly, the customary account of “reduction of the wave packet” (going back to von Neumann [2]) is formulated in terms of an observable with discrete spectrum and a corresponding ideal measurement. Secondly, in expounding the probability interpretation of quantum mechanics it is usually stated that one should think of the state vector of a physical system as containing information concerning an ensemble of identically prepared systems (i.e., described by the same state vector). Finally, a large amount of attention is traditionally paid to complete sets of commuting observables whose measurement reduces the system to a unique eigenstate of all of them. However, these notions become rather obscure if the position of the (center of mass of the) system is taken into account. Consider, for example, a free nonrelativistic particle. If we accept the

* Present address: Max-Planck-Institut für Physik und Astrophysik, Föhringer Ring 6, 8000 Munich 40, Federal Republic of Germany.
probability interpretation of the position space wavefunction, and refuse to attach a physical meaning to nonnormalizable "eigenstates" of the position operator, then even the most precise position measurements will only tell us that the particle is in a nonzero volume $V$. But the Hilbert space of wavefunctions localized in $V$ is infinite-dimensional, and there seems to be no way to ascertain that a particle may be described by a given wavefunction. One might object that this difficulty can and should be cured by using density matrices, but a moment's thought should convince one that the assignment of a density matrix is afflicted with the same arbitrariness. It is therefore not possible to prepare a state, in the usual sense of ensuring that many different particles are described by the same wavefunction or density matrix at consecutive times.

This state of affairs is in sharp contrast with the traditional accounts of quantum measurement, for example those based on the Stern-Gerlach experiment (cf., e.g., Ref. [3]). The problem may be circumvented in these cases, since one has an excellent reason (viz., the success in accounting for the observations of the remaining discrete spectrum observables) to treat the position on the same footing as the measurement apparatus. However, this does not alleviate the need for a clearcut operational meaning for the position space wave function (and more generally, wavefunctions "diagonal" w.r.t. other continuous-spectrum observables like momentum, to which similar remarks apply). In this connection, we should like to point out that the often-made assumption that the wavefunction immediately after the detection of a particle in $V$ is just the (renormalized) restriction to $V$ of the original wavefunction ("ideal position measurement") seems quite implausible in view of the behavior of particles in bubble chambers. Indeed, if one regards the various ionization events as position measurements (as is customarily done) it seems clear from the pictures one sees that the disturbance in particle momentum for each event is very small, whereas the "chopping off" just described would change energy and momentum in a very drastic fashion. For instance, a wavefunction that is smooth but for jumps at the boundary of $V$ typically has an infinite energy expectation. One might instead be tempted to argue in favor of some smooth probability density vanishing outside $V$, but it should be remembered that state vectors with the same modulus in position space may have vastly different momentum space moduli.

We shall not dwell any longer on the embarassing status of the localization problem (for which we have no cure to offer), but would like to emphasize that a better understanding would not permit the present proliferation of position operators proposed for free relativistic particles: even position operators with noncommuting and nonnormal components have been seriously considered (for a review, see Ref. [4]). Here, we will only consider the Newton-Wigner (henceforth NW) operator [5], which is the obvious relativistic analog of the uncontested nonrelativistic position operator, at least if one takes the point of view (as we do) that quantum states of a free relativistic elementary particle may be described by unit vectors in a Hilbert space carrying a unitary irreducible representation of the Poincaré group. As is well known, the mass $m$ and spin $s$ of the particle are then just the values of two Casimir operators, which serve as labels for the irreducible representation. We assume from
now on that \( m > 0 \) and \( s = 0 \). (The latter restriction is not essential; of all operators below, only the rotation and boost operators would act nontrivially on spin space if \( s > 0 \).) Then the representation can be assumed to act on the Hilbert space \( \mathcal{H}_p \) of square-integrable momentum wavefunctions,

\[
\mathcal{H}_p \equiv L^2(\mathbb{R}^3, d\vec{p}),
\]

and is given explicitly by

\[
(U(a, A)\psi)(\vec{p}) = \exp(ia \cdot p) [(A^{-1}p)_0/p_0]^{1/2} \psi(A^{-1}p).
\]

Here, \( p \) is the energy–momentum 4-vector,

\[
p = (E_p, \vec{p}), \quad E_p \equiv (|\vec{p}|^2 + m^2 c^4)^{1/2},
\]

and the dot denotes the Lorentz inner product. By applying Mackey’s imprimitivity theory to the representation of the subgroup of Euclidean motions, Wightman [6] has shown that the NW position operator

\[
\vec{x}_{NW} \equiv i\hbar \nabla_{\vec{p}}
\]

is the unique operator satisfying a number of minimal physical requirements. (More precisely, in the Heisenberg picture we employ, this is the NW operator at time zero.)

As in nonrelativistic quantum mechanics, the function

\[
\psi(x) \equiv \frac{1}{(2\pi\hbar)^{3/2}} \int d\vec{p} \exp(i\vec{x} \cdot \vec{p}/\hbar) \psi(\vec{p}),
\]

where \( \psi \in \mathcal{H}_p \) and \( \|\psi\| = 1 \), is then interpreted as a position probability amplitude. That is, the number

\[
P(V) = \int_V d\vec{x} |\psi(\vec{x})|^2
\]

is regarded as the probability of observing the particle in the spatial volume \( V \).

One of the main reasons that some authors are not satisfied with the NW position operator seems to be the so-called acausal spreading associated with it (as pointed out first by Fleming [7] and further discussed by Schlieder [8]): If the wavefunction \( \psi(x) \) vanishes outside a bounded set, then its free evolution \( (U_t\psi)(x) \) fails to have this property for any \( t \neq 0 \). (To see this (cf. Ref. [9]), note that in this case \( \psi(\vec{p}) \) is (the restriction of) an entire function, while

\[
(U_t\psi)(\vec{p}) = \exp(-itE_p/\hbar) \psi(\vec{p})
\]

is not for any \( t \neq 0 \). Indeed, if it were, the function \( \exp(-itE_p/\hbar) \) would be meromorphic as a quotient of two entire functions, which is clearly not the case in view of its branch point singularity.) Recently, it has been shown by Hegerfeldt and
the author [10] that this behavior in fact occurs for any nonconstant Hamiltonian that is a semibounded function of \( \bar{p} \); furthermore, under somewhat stronger assumptions the spreading is over all of space. It is the main purpose of the present paper to assess the extent of this behavior for the NW operator, and in particular (in Section 3) to estimate the chances of detecting acausal events in a typical laboratory situation, assuming the conventional interpretation of the wavefunction as described above. In Section 4 we discuss the result obtained in Section 3. Section 2 is less practical: We show that NW localization is, loosely speaking, asymptotically causal and introduce a function that would be the best measure of acausal behavior in a world not encumbered by experimental uncertainties and the unavailability of particles of arbitrary masses and energies. Section 5 concludes the paper with remarks on the noninvariance of the NW operator and on the velocity operator corresponding to the NW operator.

2. Asymptotic Causality

In this section and the next we consider position space wavefunctions \( \psi \) belonging to the Hilbert space

\[
\mathcal{H}_x = L^2(\mathbb{R}, dx).
\]

Also, in this section we put \( \hbar = c = 1 \). Let \( B \) be a bounded set in \( \mathbb{R}^3 \), and define

\[
B_t = \{ x' \in \mathbb{R}^3 \mid \text{dist}(x', B) \leq t \}.
\]

Thus, \( B_t \) is the set of points that can be reached by light signals from \( B \) in time \( t \). Put, for \( V \subset \mathbb{R}^3 \),

\[
\chi_V(x) = 1, \quad x \in V,
\]

\[
= 0, \quad x \notin V.
\]

Then the quantity

\[
P_t = \|(1 - \chi_{B_t}) U_t \psi \|^2, \quad \text{supp} \ \psi \subset B, \ |\psi| = 1,
\]

where we have used the notation \( |\psi|^2 = \int dx |\psi(x)|^2 \), where \( \chi_V \) denotes multiplication by \( \chi_V(x) \), and where \( U_t \) is the free evolution on \( \mathcal{H}_x \), can be interpreted as the probability of observing the wave packet \( \psi \) supported in \( B \) at time zero outside the domain of dependence of \( B \) at time \( t \). We claim that for any such \( \psi \) one has

\[
\lim_{t \to \infty} P_t = 0.
\]

Thus, NW localization is "asymptotically causal." This claim clearly follows from the following much stronger assertion.
LOCALIZATION AND SUPERLUMINAL SPEEDS

PROPOSITION. Let $\psi \in \mathcal{H}$ and $\tilde{x}_n \in \mathbb{R}^3$. Then

$$\lim_{t \to \infty} \int_{|x - \tilde{x}_n| > t} d\tilde{x} |(U_t \psi)(\tilde{x})|^2 = 0.$$  \hspace{1cm} (2.6)

Proof. By making a change of variables we may assume $\tilde{x}_n = 0$. Also, since $U_t$ is unitary, we need only show that (2.6) holds for a set $\mathcal{D}$ of wavefunctions that is dense in $\mathcal{H}$. (That is, for any $\psi \in \mathcal{H}$ one can find $\psi_n \in \mathcal{D}$ such that $\|\psi - \psi_n\| < 1/n$.) We take for $\mathcal{D}$ the dense set of wavefunctions $\psi$ whose Fourier transforms $f$ are infinitely differentiable and have bounded support. Then we may write

$$(U_t \psi)(\tilde{x}) = \int d\tilde{p} f(\tilde{p}) \left( \frac{(\tilde{x} + \tilde{p}E_p^{-1}t) \cdot \nabla_{\tilde{p}}}{r^2 - |\tilde{p}|^2 E_p^{-2}t^2} \right)^2 e^{i\tilde{p} \cdot \tilde{x} - it\tilde{p}},$$  \hspace{1cm} (2.7)

where $r = |\tilde{x}|$ and $f \in C_0^\infty(\mathbb{R}^3)$. But since $m > 0$ we have

$$r^2 - |\tilde{p}|^2 E_p^{-2}t^2 \geq Cr^2, \quad \forall r \geq t, \; \forall \tilde{p} \in \text{supp} f,$$

where $C > 0$, so for $r \geq t$ we may integrate by parts and estimate in the obvious way, obtaining

$$|(U_t \psi)(\tilde{x})| \leq C' r^{-2}, \quad \forall r \geq t.$$  \hspace{1cm} (2.9)

Hence,

$$\int_{r \geq t} d\tilde{x} |(U_t \psi)(\tilde{x})|^2 \leq C'' \int_t^\infty \frac{dr}{r^2} = \frac{C''}{t}.$$  \hspace{1cm} (2.10)

Since the right-hand side vanishes for $t \to \infty$, (2.6) follows. Q.E.D.

Note that the proof makes use of the assumption that $m$ is strictly positive. This restriction is essential, since the corresponding assertion is false if $m = 0$, as has been shown by Petzold [11]. Notice also that this proposition implies that (2.5) actually holds for any $\psi \in \mathcal{H}$. Thus, this property contains no special information about the states that are originally localized in $B$. In particular, the above does not show how fast the probability vanishes in general for such states, and how large it can be. As an appropriate measure for this, one might take the supremum of $P_t$ over all $\psi \in \mathcal{H}$ with $\text{supp} \psi \subset B$ and $\|\psi\| = 1$. For simplicity, let us take for $B$ the ball $B_R$ of radius $R$ with center at the origin, and let us put $\chi_R = \chi_{B_R}$. Then this supremum can be written in terms of an operator norm as

$$P(t, m, R) \equiv \| (1 - \chi_{B_R}) U_t \chi_R \|^2,$$  \hspace{1cm} (2.11)

where, as before, $U_t$ is the pseudo-differential operator $\exp[-i(-A + m^2)^{1/2} t]$ (i.e., the Fourier transform of multiplication by $\exp(-iE_p t)$).
Unfortunately, we were unable to evaluate $P(t, m, R)$ or to obtain some illuminating bound on it. (The relation

$$P(t, m, R) = P(t/\lambda, m\lambda, R/\lambda), \quad \forall \lambda \in (0, \infty), \quad (2.12)$$

which follows from scaling, may possibly be helpful in this regard.) We cannot even exclude the possibility that $P(t, m, R) = 1$ for $t \neq 0$. This would mean that for any given $t \neq 0$ one could find states localized in $B_R$ at $t = 0$ with a probability of being outside $B_{R+t}$ at time $t$ that is arbitrarily close to 1. However, by analyzing this problem in a more physical way in the next section, we shall reach the conclusion that even if this were so, the prospects of observing such states are very dim indeed.

3. Can Superluminal Propagation be seen?

An experiment to detect acausal events might be set up along the following lines. Surround a source of particles with a spherical particle detector of outer radius $R$ and surround this detector with a concentric spherical detector of inner radius $R + d$. Ensure vacuum conditions between the two detectors. Register the time $t_l$ when a particle leaves the inner detector and the time $t_e$ when it enters the outer one. (Of course, the particle flux must be sufficiently low for this to be possible.) If

$$t_e - t_l \equiv T < d/c, \quad (3.1)$$

the particle has apparently crossed the gap at a superluminal speed, so such events would be counted as acausal.

Let us therefore introduce a function $P(d, \gamma)$ denoting the probability that the particle will enter the outer detector in the time interval $(0, \gamma d/c)$ (where $0 < \gamma < 1$), given that it left the inner one at $t = 0$. Assuming as before that the localization of the particle is described by the NW operator, we shall now derive an upper bound on $P(d, \gamma)$ and use this to estimate the probability for observing acausal events, taking into account experimental uncertainties and the unavailability of easily detectable massive particles lighter than the electron. (The reader may wish to skip the derivation at first reading and proceed to the result, Eq. (3.10).)

We first observe that

$$P(d, \gamma) \leq F(d, \gamma d/c)^2, \quad (3.2)$$

where

$$F(d, T) \equiv \|(1 - \chi_{R+d}) U_T \chi_R\|. \quad (3.3)$$

(This is actually not as obvious as it seems, cf. Section 4.) Thus, we need only
estimate $F(d, T)$ for $T < d/c$. Since the free evolution acts as multiplication by $\exp(-itE_p/\hbar)$ on $\mathcal{F}_p$, $U_T$ acts on $\mathcal{F}_p$ as a convolution operator with kernel

$$C_T(\vec{x}) = \lim_{\epsilon \to 0} (2\pi\hbar)^{-3} \int d\vec{p} \exp(i\hbar^{-1}|p \cdot \vec{x} - E_p(T - i\epsilon)|)$$

$$- \lim_{\epsilon \to 0} \frac{-1}{2\pi^2 r} \partial_r \int_0^\infty dq \cos qr \exp[-i(|q|^2 + \lambda^{-2})c(T - i\epsilon)], \quad r > CT, \quad (3.4)$$

were we have introduced the Compton wavelength,

$$\lambda = \hbar/mc. \quad (3.5)$$

The integral occurring here is known (cf. Ref. [12], p. 16, Eq. (26)), and using this we obtain

$$C_T(\vec{x}) = -\frac{icT}{2\pi r\lambda} \partial_r K_1(\lambda^{-1}|r^2 - c^2T^2|^{1/2})(r^2 - c^2T^2)^{-1/2}, \quad r > cT, \quad (3.6)$$

where $K_1$ is a modified Bessel function. We now note that

$$\| (1 - \chi_{R+d}) U_T \chi_R \psi \|^2$$

$$= \int_{r > R+d} d\vec{x} \left| \int_{r' < R} d\vec{x}' C_T(\vec{x} - \vec{x}') \psi(\vec{x}') \right|^2$$

$$= \int_{r > R+d} d\vec{x} \left| \int_{r' < R} d\vec{x}' [1 - \chi_d(\vec{x} - \vec{x}')] C_T(\vec{x} - \vec{x}') \psi(\vec{x}') \right|^2$$

$$\leq \int d\vec{x} \left| \int d\vec{x}' [1 - \chi_d(\vec{x} - \vec{x}')] C_T(\vec{x} - \vec{x}') \psi(\vec{x}') \right|^2. \quad (3.7)$$

Thus, $F(d, T)$ is bounded above by the norm of the operator of convolution with the positive function $|1 - \chi_d(\vec{x})| C_T(\vec{x})$. But as is well known (and easily seen) the norm of such an operator equals the $L^1$-norm of the function, so that

$$F(d, T) \leqslant -i \int_{r > d} d\vec{x} C_T(\vec{x})$$

$$= -\frac{2cT}{\pi\lambda} \int_0^\infty dr \partial_r [K_1(\lambda^{-1}|r^2 - c^2T^2|^{1/2})(r^2 - c^2T^2)^{-1/2}]$$

$$= \frac{2cdTd}{\pi\lambda} K_1(\lambda^{-1}|d^2 - c^2T^2|^{1/2})(d^2 - c^2T^2)^{-1/2}$$

$$+ \frac{2cT}{\pi\lambda} \int_0^\infty dr K_1(\lambda^{-1}|r^2 - c^2T^2|^{1/2})(r^2 - c^2T^2)^{-1/2}$$
\[ \frac{2cTd}{\pi \lambda} K_1(\lambda^{-1}[d^2 - c^2 T^2]^{1/2})(d^2 - c^2 T^2)^{-1/2} + \frac{2cT}{\pi d} K_0(\lambda^{-1}[d^2 - c^2 T^2]^{1/2}), \]  

(3.8)

where we used integration by parts and the relation \( K_1(x) = -K_0'(x) \). Now one has, for \( x > 0 \),

\[ K_0(x) = \int_0^\infty d\theta \exp(-x \cosh \theta) < K_1(x) \]

\[ = \int_0^\infty d\theta \cosh \theta \exp(-x \cosh \theta) \]

\[ \leq \int_0^\infty d\theta e^\theta \exp\left(-\frac{1}{2}xe^\theta\right) = \left(\frac{2}{x}\right) e^{-x/2}. \]  

(3.9)

Hence, using (3.2) and (3.8) we finally get the desired bound:

\[ P(d, \gamma) < \frac{16\gamma^2}{\pi^2(1 - \gamma^2)^2} \left(1 - \gamma^2\right)^{-1/2} + \frac{\lambda}{d} \right)^2 \exp \left( -\frac{d}{\lambda} \left(1 - \gamma^2\right)^{1/2} \right), \]  

(3.10)

where \( \lambda \) is the Compton wavelength (3.5) and

\[ \gamma = cT/d. \]  

(3.11)

We now note that the measurements of \( T \) and \( d \) and the value of \( c \) are afflicted with certain experimental uncertainties. Let us denote the relative error in \( \gamma \) by \( \varepsilon \). For an event to be counted as acausal, its corresponding \( \gamma \) should be smaller than \( 1 - \varepsilon \). Since the right-hand side of (3.10) is an increasing function of \( \gamma \) we may therefore estimate the probability of superluminal events by taking \( \gamma = 1 - \varepsilon \). If we now assume that the absolute error in \( d \) and \( T \) is independent of \( d \) (which is reasonable for distances in a laboratory), then \( \varepsilon \) is inversely proportional to \( d \) in the regime where \( c \) may be considered to be known precisely (i.e., as long as \( \varepsilon \gg 4 \times 10^{-9} \)). Hence, to maximize the right-hand side as a function of \( d \), we should take \( d \) as small as possible. However, it is clear that one should not make \( d \) smaller than the distance that light travels in the resolution time of the detection devices. Let us assume therefore (somewhat arbitrarily) that \( d = 1.00 \times 10^{-3} \) m and \( \varepsilon = 10^{-2} \). This constitutes a generous estimate, since the corresponding resolution time is well beyond current technology. Also, to maximize the right-hand side of (3.10) as a function of \( \lambda \), we should take \( \lambda \) as big as possible. Thus, one should perform the experiment with electrons, for which \( \lambda \approx 4 \times 10^{-13} \) m. Estimating the right-hand side for \( \gamma = 1 - 10^{-2}, \lambda = 4 \times 10^{-13} \) m and \( d = 10^{-3} \) m, we finally conclude that the probability of detecting a superluminal electron in this experiment is smaller than \( 10^{-18} \).
We leave it to the reader to answer the question that heads this section. He is reminded that we assumed a resolution time on the order of 0.03 psec, and that the lifetime of the universe is believed to be smaller than $10^{30}$ psec.

4. Discussion

Let us now comment in more detail on the assumptions and results of Section 3. Firstly, as long as the particle is observed by the inner detector, its wavefunction is assumed to satisfy $\psi = x_R \psi$ according to the standard interpretation of quantum mechanics. The expression $\| (1 - x_{R+d}) U_T \psi \|^2$ is then interpreted as the probability that the answer to the question: “After a time $T$, is the particle located outside the sphere of radius $R + d$ centered at the origin?” is “yes.” What is understood here, is that this yes–no experiment is done at time $T$ and that the wave packet has evolved freely until then. This is, however, not how an actual experiment would be done; the detector is set up beforehand and may be triggered before time $T$. The corresponding “reduction of the wave packet” implies that such a particle did not evolve freely until time $T$. (One might even take the point of view that the wave packet is continually being reduced in view of the presence of the detectors!) Unfortunately, the conventional approach to quantum mechanics does not seem to consider such situations and, in the case at hand, does not provide an expression for the probability of detection in the time interval $(0, T]$. We have therefore made the assumption that the latter probability for the finite outer detector is bounded above by the former probability for the infinite outer region. Since a particle that has “passed through” the outer detector before $T$ “is” in the outer region (classically speaking), this assumption is highly plausible, but it is not completely cogent.

Secondly, we have ignored the problem of what kind of states would be most likely to exhibit superluminal propagation. This we could do, since our bounds hold true for any state, being formulated in terms of operator norms. The fact that the particle source can only supply particles of finite and limited energies will presumably reduce the probability of observing acausal events by another staggering factor.

Thirdly, one may object that theoretical arguments should not pay attention to what is experimentally feasible, but should rather concentrate on what is in principle feasible. Surely, one might argue, the bound (3.10) does not exclude a noticeable probability for acausal events if one could vary $d$ and $\lambda$ at will, and if one could make experimental errors arbitrarily small. To this we should like to remark that the unpalatability of the “acausality” under consideration here mainly depends on one’s a priori conceptions. It seems to us that the only legitimate objection to the superluminal events predicted by the NW operator is that they have not been observed. This is, however, completely in agreement with the above result, which shows that to all intents and purposes such events are (presently) unobservable.
5. Concluding Remarks

(i) Noninvariance of the NW Operator

Another important reason why the NW operator is often considered unsatisfactory is its so-called noninvariance under Lorentz boosts. This is usually formulated in terms of "eigenfunctions" localized at one point in one inertial frame. However, since nonnormalizable functions are not covered by the conventional interpretation of quantum mechanics, we prefer to look at this in the following way: If a wavefunction \( \psi(\vec{x}) \in H_x \) vanishes outside a bounded set, this is in general false for \( (U(A)\psi)(\vec{x}) \) if \( A \) is a boost, since the (extension of the) Fourier transform of the latter wavefunction has branch point singularities in general (cf. Eq. (1.2)). (In fact, it is plausible this is always the case.) It would be of interest to investigate the extent of this boost delocalization along the same lines as we did above for the time evolution. However, no matter what the result of such an analysis might be, in this case one would be hard put to propose a realistic experiment that would "measure" this phenomenon. Indeed, the usual invariance requirements only concern experiments that are done in one fixed inertial frame and their transcription to other frames; they do not concern measurements of the same property in different frames, for which the concept of simultaneity loses its meaning. Remember also that even at the classical level space-time symmetry principles do not imply that qualitative properties are invariant under Lorentz boosts; one can think for instance of the vanishing magnetic field of a stationary charge and the isotropy of the microwave background. Thus, we feel it is largely a matter of choice whether one is willing to describe a particle with a wavefunction that is localized only in the special frames in which the detection device is at rest.

(ii) A Paradox

The velocity operator corresponding to the NW operator is given by

\[ \vec{v}_{NW} = \hbar^{-1}i[H, \vec{x}_{NW}] = \vec{p}c^2/E_p \quad (5.1) \]

and therefore predicts propagation speeds that are always smaller than \( c \). On the other hand, because of the instantaneous spreading of the wave packet there is a nonzero (albeit practically negligible) probability to measure arbitrarily large speeds. To us this inconsistency is more unsettling than the acausality and noninvariance associated with the NW operator. It is, in particular, not clear to us whether it should be considered as an argument against the NW operator, as an indication that its time derivative should only be regarded as a velocity operator in the average sense of Ehrenfest's theorem, or as a hint that its physical interpretation needs revision.

Acknowledgments

We would like to thank G. Battle and F. Narcowich for critical comments and suggestions, and R. Kenefick and R. Webb for discussions on experimental aspects.
REFERENCES