
To cite this article: Iwo Bialynicki-Birula and Zofia Bialynicka-Birula 2019 New J. Phys. 21 119502

View the article online for updates and enhancements.
In the equations (19) and (23) the mixed terms containing the products of the amplitudes \( f(p, +) \) and \( f(p, -) \) are missing. There was also an additional misprint.

The corrected form of equation (19) is:

\[
\Delta r^2 = \frac{1}{N^2} \int d^3p \left[ |\nabla_p f(p, s)|^2 + \left( \frac{1}{p^2} - \frac{m}{p^2 E_p} + \frac{m^2}{4E_p^2} \right) f(p, s)^2 \right]
\]

\[
- \frac{i}{2\pi} \left( 1 - \frac{m}{E_p} \right) p_\perp f^*(p, s) \partial_\perp f(p, s) - p_\perp f^*(p, s) \partial_\perp f(p, s)
\]

\[
+ \frac{1}{2N^2} \int d^3p \left( 1 - \frac{m}{E_p} \right) |(p_\perp - i\vec{p}_\perp) f^*(p, +) \partial_\perp f(p, -) - p_\perp f^*(p, +) (\partial_\perp - i\vec{\partial}_\perp) f(p, -) + c.c. \right]
\]

where \( \partial_\perp = \partial / \partial_{R_\perp} \), \( \partial_\parallel = \partial / \partial_{R_\parallel} \), \( \partial_\parallel = \partial / \partial_{R_\parallel} \).

The corrected form of equation (23) is:

\[
\Delta r^2 = \frac{1}{N^2} \int_0^\infty dp \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \left[ p^2 |\partial_\phi f(p, \theta, \phi, s)|^2 + |\csc \theta \partial_\theta f(p, \theta, \phi, s)|^2 \right]
\]

\[
+ \left( 1 - \frac{m}{E_p} + \frac{m^2}{4E_p^2} \right) f(p, \theta, \phi, s)^2 - \frac{i}{2} \left( 1 - \frac{m}{E_p} \right) f^*(p, \theta, \phi, s) \partial_\theta f(p, \theta, \phi, s)
\]

\[
+ \frac{1}{2N^2} \int_0^\infty dp \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \left[ 1 - \frac{m}{E_p} \right]
\]

\[
\times \left( f^*(p, \theta, \phi, +) \partial_\theta f(p, \theta, \phi, -) + i \cot \theta f^*(p, \theta, \phi, +) \partial_\phi f(p, \theta, \phi, -) \right) e^{-i\phi} + c.c. \right]
\]

These changes do not influence our final conclusion since all omitted terms contain the derivatives with respect to angular variables while the functions that minimize the product of uncertainties are spherically symmetric.
Heisenberg uncertainty relation for relativistic electrons

To cite this article: Iwo Bialynicki-Birula and Zofia Bialynicka-Birula 2019 New J. Phys. 21 073036

View the article online for updates and enhancements.
Heisenberg uncertainty relation for relativistic electrons

Iwo Bialynicki-Birula and Zofia Bialynicka-Birula

1 Center for Theoretical Physics, Polish Academy of Sciences Aleja Lotników 32/46, 02-668 Warsaw, Poland
2 Institute of Physics, Polish Academy of Sciences Aleja Lotników 32/46, 02-668 Warsaw, Poland

E-mail: birula@cft.edu.pl

Keywords: uncertainty relations, relativistic electrons, Dirac equation

Abstract

The Heisenberg uncertainty relation is derived for relativistic electrons described by the Dirac equation. The standard nonrelativistic lower bound $\frac{3}{2} \hbar$ is attained only in the limit and the wave function that reproduces this value is singular. At the other end, in the ultrarelativistic limit, the bound is the same as that found before for photons.

1. Introduction

Recent progress in the generation of relativistic beams of electrons calls for adequate theoretical tools, based on the Dirac equation, that would allow to account precisely for quantum properties of these electrons. The problem raised and answered in this paper is that of Heisenberg uncertainty relation (HUR) for relativistic electrons. It is shown that the relativistic HUR is markedly different from its original nonrelativistic form.

In most monographs [1–8] devoted to relativistic quantum mechanics of electrons the HURs has not been mentioned. In those few places where there is a reference to uncertainty relations in a relativistic theory, the authors [9, 10] follow the early paper by Landau and Peierls [11] giving only a qualitative analysis of limitations imposed by relativity on measurements. In this work we present a complementary approach to uncertainty relations. It could be said that we do what Kennard [12] had done for the nonrelativistic form of the uncertainty relation. He derived an exact mathematical inequality which expressed precisely the heuristic explanation (Bohr–Heisenberg microscope) of the nonrelativistic uncertainty relation. We do the same for relativistic electrons.

In nonrelativistic quantum mechanics one may rely directly on the interpretation of the modulus squared of the wave function $|\psi(r, t)|^2$ and its Fourier transform $|\tilde{\psi}(p, t)|^2$ as the probability densities in position and momentum space. In relativistic quantum mechanics there is an ambiguity in the definition of the probability density in position space because there are various properties of relativistic particles that may be used to define their position. In particular, we may use the energy distribution, as we have done for photons [13], or the rest mass distribution.

In this work we follow Dirac [3] who postulated (p 258) that 'The square of the modulus of the wave function, summed over the four components, should give the probability per unit volume of the electron being at a certain place'. This quantity is also ([3] p 260) the time component $j^0$ of the electromagnetic current density $j^\mu$. The continuity equation $\partial_0 j^0 = 0$ guarantees the conservation of the total probability,

$$\int d^3r |\rho_0(r, t)| = \text{const},$$

(1)

where

$$\rho_0(r, t) = |\psi(r, t)|^2 = \sum_\alpha |\psi_\alpha(r, t)|^2,$$

(2)

and $\psi_\alpha(r, t)$ are the components of the Dirac bispinor. Thus, the localization of the electron is determined by its interaction with the electromagnetic field: the electron is where its charge is. A natural definition of the probability density in momentum space is the modulus squared of the Fourier transform of $\psi(r)$,
\[ \rho_p(p, t) = |\tilde{\psi}(p, t)|^2 = \sum_{\alpha} |\tilde{\psi}_{\alpha}(p, t)|^2, \]  

where

\[ \tilde{\psi}_{\alpha}(p, t) = \int \frac{d^3r}{(2\pi)^{3/2}} \psi_{\alpha}(r, t) e^{-ipr/\hbar}. \]  

Thus, the relation between \( \rho_r(r, t) \) and \( \rho_p(p, t) \), based on the Fourier transformation, is the same as in the nonrelativistic quantum mechanics.

From the probability densities in position and in momentum space we construct the standard expressions for the dispersions \( \Delta r^2 \) and \( \Delta p^2 \):

\[ \Delta r^2 = \frac{1}{N^2} \int d^3r (r - \langle r \rangle)^2 \rho_r(r), \]  
\[ \Delta p^2 = \frac{1}{N^2} \int d^3p (p - \langle p \rangle)^2 \rho_p(p), \]

where \( N^2 \) is the normalization constant

\[ N^2 = \int d^3r |\psi(r)|^2 = \int d^3p |\tilde{\psi}(p)|^2. \]

We omitted the time dependence because the uncertainty relations are always formulated at a fixed time.

A simple textbook proof of the HUR in nonrelativistic quantum mechanics \([14, 15]\) can be extended to the relativistic theory. This extension starts from the following non-negative expression built from the Dirac bispinor:

\[ \int d^3r (r - \langle r \rangle)^2 + i(\hbar \nabla - \langle p \rangle) \psi^2 \geq 0. \]

After the rearrangement, this inequality becomes a quadratic form in \( \zeta \):

\[ Q(\zeta) = \Delta r^2 + \zeta^2 \Delta p^2 + 3\zeta \hbar \geq 0, \]

where we used the expression for \( \Delta p^2 \) in the position representation

\[ \Delta p^2 = \frac{1}{N^2} \int d^3r \left( \frac{\hbar}{i} \nabla - \langle p \rangle \right) \psi^2. \]

The requirement that \( Q(\zeta) \) has at most one real zero (the discriminant cannot be positive) leads to the standard nonrelativistic HUR in three-dimensions

\[ \sqrt{\Delta r^2 \Delta p^2} \geq \frac{3}{2} \hbar. \]

Our task is to determine whether in the relativistic theory this inequality is saturated or whether it is a strict inequality. Saturation would mean that there exists a bispinor describing an electronic state which obeys the set of three equations

\[ \left( r - \langle r \rangle \right) - i\zeta \left( \frac{\hbar}{i} \nabla - \langle p \rangle \right) \psi = 0. \]

In the nonrelativistic quantum mechanics, these equations are satisfied by a Gaussian wave function but in the relativistic case it is not so simple. Of course, a Gaussian wave function is a solution the equation (12) but such a wave function does not describe a state of an electron but it will also have a part describing the positron. We shall analyze this problem in detail in the next section with the use of a variational procedure.

### 2. Evaluation of the variational functional

The variational procedure is simplified if we put \( \langle r \rangle = 0 \) and \( \langle p \rangle = 0 \) in the definitions (5) and (6). In the case of \( \Delta r^2 \) the prescription is simple. We just choose the origin of the coordinate system at the center of the probability distribution. In the case of \( \Delta p^2 \) the solution is more subtle because the space of momenta is not homogeneous and we cannot choose its origin at will. We may, however, make the following replacement:

\[ \psi(r) \rightarrow \exp i \left( \langle p \rangle \cdot r \right) \psi(r) \]

which does not restrict the freedom in the choice of the variational trial function. Such a replacement does not change \( \Delta r^2 \) but it eliminates \( \langle p \rangle \) in (10), so that we obtain:

\[ \Delta p^2 = \frac{1}{N^2} \int d^3p P^2 \rho_p(p). \]

In this way we can eliminate completely the average values of position and momentum from the formulas for the dispersions.
We will express now the product \( \Delta r^2 \Delta p^2 \) in terms of the independent degrees of freedom that can be subjected to unrestricted variations. The solution of the differential equation obtained from the variational procedure will determine the minimal value of \( \sqrt{\Delta r^2 \Delta p^2} \).

The states of an electron are described by the wave functions obeying the Dirac equation (\( \varepsilon = 1, \hbar = 1 \)),

\[
[i\gamma^\mu \partial_\mu - m] \psi(r, t) = 0.
\]

(14)

Since we are interested in the wave functions of an electron and not of a positron, we consider only positive energy solutions of the Dirac equation. Such bispinors \( \psi_\alpha(r, t) \) can be expressed in terms of two complex amplitudes \( f(p, \pm) \):

\[
\psi_\alpha(r, t) = \int \frac{dp}{(2\pi)^{3/2}} \sum_{s=\pm} u_\alpha(p, s) f(p, s) \exp(i p \cdot r - iE_p t).
\]

(15)

where \( E_p = \sqrt{m^2 + p^2} \).

The wave functions \( f(p, s) \) of momentum variables and the spin index \( s \) are not subjected to any restrictions; they represent the independent degrees of freedom of an electron moving in free space. They form a representation of the Poincaré group (inhomogeneous Lorentz group) for massive spin 1/2 particles, as has been described by Wigner [16].

Two orthonormal bispinors \( u_\alpha(p, s) \) in the Weyl (chiral) representation of \( \gamma \) matrices [17] will be chosen in the form:

\[
u(p, +) = \frac{1}{\sqrt{4E_p(E_p + m)}} \begin{bmatrix} m + E_p + p_z \\ p_x + i p_y \\ -p_x - i p_y \\ m + E_p + p_z \end{bmatrix},
\]

(16a)

\[
u(p, -) = \frac{1}{\sqrt{4E_p(E_p + m)}} \begin{bmatrix} p_x - i p_y \\ -p_x + i p_y \\ m + E_p - p_z \\ m + E_p + p_z \end{bmatrix}.
\]

(16b)

They are linear combinations of the simpler bispinors used before [18] but they are more convenient in setting up the variational functional. The Weyl representation of \( \gamma^\mu \) matrices is better suited for the relativistic analysis than the commonly used Dirac representation because the upper and lower components are then genuine relativistic spinors.

To set up the stage for the variational calculation, we shall express the dispersions in terms the amplitudes \( f(p, s) \). Orthonormality of the bispinors \( u(p, s) \) leads immediately to simple forms of and \( N^2 \) and \( \Delta p^2 \),

\[
N^2 = \int dp \sum_s |f(p, s)|^2,
\]

(17)

\[
\Delta p^2 = \frac{1}{N^2} \int dp \sum_s |f(p, s)|^2.
\]

(18)

The corresponding formula for \( \Delta r^2 \),

\[
\Delta r^2 = \frac{1}{N^2} \int dp \sum_s \left[ |\nabla_p f(p, s)|^2 + \left( 1 - \frac{m}{E_p} + \frac{m^2 p^2}{4E_p^2} \right) |f(p, s)|^2 \right] - \frac{s}{2} \left( 1 - \frac{m}{E_p} \right) \left( p_x f^*(p, s) \frac{\partial}{\partial p_x} f(p, s) - p_y f^*(p, s) \frac{\partial}{\partial p_y} f(p, s) \right)
\]

(19)

is obtained with the use of the following identity:

\[
r \psi_\alpha(r) = \int \frac{dp}{(2\pi)^{3/2}} i \nabla_p \left( \sum_s u_\alpha(p, s) f(p, s) \right) \exp(i p \cdot r).
\]

(20)

Double arrows in (19) denote the antisymmetrized derivative, i.e. \( \overset{\leftrightarrow}{\partial \phi} = \partial \phi - g \partial \phi f \). We now express these formulas in spherical coordinates to simplify the variational equations.

\[
N^2 = \int_0^\infty dp \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \sum_s |f(p, \theta, \phi, s)|^2,
\]

(21)
\[
\Delta p^2 = \frac{1}{N^2} \int_0^\infty dp \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \sum_i p^2 |f(p, \theta, \phi, s)|^2, 
\]

(22)

\[
\Delta r^2 = \frac{1}{N^2} \int_0^\infty dp \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \sum_i \left[ p^4 \partial_p f(p, \theta, \phi, s) \right]^2 + \left[ \partial_p f(p, \theta, \phi, s) \right]^2 + \left[ |\csc \theta \partial_\phi f(p, \theta, \phi, s)| \right]^2 
- \frac{i s}{2} \left[ \frac{m}{E_p} \right] f^*(p, \theta, \phi, s) \partial_p f(p, \theta, \phi, s) + \left( 1 - \frac{m}{E_p} + \frac{m^2 + p^2}{4E_p^2} \right) |f(p, \theta, \phi, s)|^2 \right]. 
\]

(23)

The product \( \sqrt{\Delta r^2 \Delta p^2} \) of the dispersions is a well defined functional of the electron amplitudes \( f(p, \theta, \phi, s) \). The minimal value of this product that appears on the right-hand side of the uncertainty relation (11) will be found in the next section.

3. Solutions of the variational equations

We will search for the function that minimizes the uncertainty relation only among spherically symmetric functions \( f(p) \). In this case the variation of \( \Delta r^2 \Delta p^2 \) with respect to \( f^* \) gives the following ordinary differential equation for \( f \):

\[
\left[ \Delta p^2 \left\{ -\frac{\partial_p^2}{p} - \frac{2}{p} \partial_p + \frac{1}{p^2} \right\} - \frac{m}{p^2 \sqrt{m^2 + p^2}} + \frac{m^2}{4(m^2 + p^2)^2} \right] + \Delta r^2 p^2 - 2 \gamma^2 \right] f(p) = 0, 
\]

(24)

where \( \gamma = \sqrt{\Delta r^2 \Delta p^2 / \hbar} \). We dropped the index \( s \) because the equations for both components of \( f(p, s) \) are identical. In terms of the dimensionless momentum variable \( q \) and the dimensionless parameter \( d \), the variational equation takes on the following form:

\[
\left[ \frac{1}{2} \left\{ -\partial_q^2 - \frac{2}{q} \partial_q + \frac{1}{q^2} \right\} - \frac{1}{q^2 \sqrt{1 + d^2 q^2}} + \frac{d^2}{4(1 + d^2 q^2)^2} + q^2 \right] f_d(q) = \gamma f_d(q), 
\]

(25)

where

\[
q = \frac{p}{mc d}, \quad d = \left( \frac{\hbar^2 \Delta p^2}{\Delta r^2} \right)^{1/4}, 
\]

(26)

and we explicitly indicated the dependence of the function \( f \) on the parameter \( d \).

Thus, our variational equation has the form of the eigenvalue problem for the radial wave equation with a modified harmonic oscillator potential. The lowest value of \( \gamma \) (the energy of the ground state) will be the best bound in the uncertainty relation. Should we allow for the angular dependence, the effective potential would pick up the centrifugal term \( \hbar l (l + 1) / q^2 \). This would definitely increase the ground state energy.

The eigenvalue problem (25) has no analytic solution but considering the potential \( V(q) \),

\[
V(q) = \frac{1}{q^2} - \frac{1}{q^2 \sqrt{1 + d^2 q^2}} + \frac{d^2}{4(1 + d^2 q^2)^2} + q^2, 
\]

(27)

as a function of the parameter \( d \) we find (see figure 1) that the smaller the value of \( d \), the deeper the potential. Therefore, we will obtain lower values of the eigenvalue \( \gamma \) for smaller values of \( d \). In the formal nonrelativistic limit (when \( m \to \infty \) or \( c \to \infty \)) we have \( d \to 0 \) and the variational equation is that of the harmonic oscillator in 3D

\[
\left[ \frac{1}{2} \left\{ -\partial_q^2 - \frac{2}{q} \partial_q + q^2 \right\} \right] f_0(q) = \gamma_0 f_0(q). 
\]

(28)

Therefore, the wave function \( f_0(q) \) is a Gaussian,

\[
f_0(q) = e^{-\gamma_0 q^2}, 
\]

(29)

and we obtain the nonrelativistic result \( \gamma_0 = 3/2 \). The physical limit of \( d \to 0 \) is achieved when either the dispersion in \( p \) tends to 0 or the dispersion in \( r \) tends to infinity. Of course, these limits cannot be reached for regular, integrable functions. Thus, in relativistic quantum mechanics of electrons the HUR has the form of the strict inequality:

\[
\sqrt{\Delta r^2 \Delta p^2} > \frac{3}{2} \hbar. 
\]

(30)
In contrast to the nonrelativistic case, this inequality is never saturated. There is no square integrable function that would give the equality in (30), but one may get arbitrarily close to $\frac{3}{2}$.

An exact form of the relativistic uncertainty relation can be obtained also in the other extreme case; in the ultra-relativistic limit, when $d \to \infty$. In this case, the variational equation

$$\frac{1}{2} \left[ -\partial_q^2 - \frac{2}{q} \partial_q + \frac{1}{q^2} + q^2 \right] f_\infty(q) = \gamma_\infty f_\infty(q)$$

again has an analytic solution. The wave function of the lowest energy state and the corresponding value of $\gamma$ are:

$$f_\infty(q) = q^{(\sqrt{5}-1)/2} e^{-q^2/2}, \quad \gamma_\infty = 3/2 \sqrt{1 + \frac{4\sqrt{5}}{9}} = 1 + \sqrt{5}/2. \quad (32)$$

Quite unexpectedly, this value of $\gamma_\infty$ is the same as for photons [13], even though the definition of $\Delta r^2$ for photons is different; it is based on the energy density and not on the charge density.

4. Two explicit examples

The general results presented here are well illustrated by two analytic solutions of the Dirac equation: Ground state of the electron in the Coulomb potential and the Dirac hopping (discovered by us recently [19]).

4.1. Hydrogen-like ion

The Dirac bispinor describing the ground state of the electron in a hydrogen-like ion has the form ([5] p 55):

$$\Psi_0(r, \theta, \phi, t) = N_G r^{\gamma-1} e^{-\sqrt{1-\gamma^2}r} e^{-i\Phi t} \begin{pmatrix} 1 \\ 0 \\ i \frac{1-\gamma}{\sqrt{1+\gamma}} \cos \theta \\ -i \frac{1-\gamma}{\sqrt{1+\gamma}} e^{i\Phi} \sin \theta \end{pmatrix}$$

where $\gamma = \sqrt{1 - \alpha^2 Z^2}$, the unit of length is the electron Compton wave length $\hbar/mc$, and the normalization constant $N_G$ is:

$$N_G = \frac{2^{\gamma} (1 + \gamma)^{\gamma+1/2} (1 - \gamma)^{\gamma+1/2}}{4\pi \Gamma(1 + 2\gamma)}. \quad (34)$$

We use the standard notation for the hydrogenic wave functions, hoping that the use of $\gamma$ in this subsection will not be confused with its use in the uncertainty relations. For the bispinor (33) both dispersions $\Delta r^2$ and $\Delta p^2$ can be analytically evaluated and the HUR reads:

![Figure 1. Dependence of the potential $V(q)$ on the parameter $d$. The horizontal lines mark the values of the five ground state energies $\gamma$ of the eigenvalue problem (25) obtained for the same values of $d$. The eigenvalues for the limiting cases $d = 0$ and $d = \infty$ are represented by dashed lines.](image-url)
Our parameter $d$ which characterizes the transition from the nonrelativistic to the highly relativistic regime has the form:

$$d = \left( \frac{2(1 + \gamma)(1 - \gamma)^2(2 - \gamma)}{\gamma(4\gamma^2 - 1)} \right)^{1/4}. \tag{35}$$

For hydrogen, when the Coulomb field is weak ($\gamma \approx 1, d \approx 0$) the rhs in (35) is $\hbar\sqrt{7/2}$, so it is still much higher than the lower bound $3/2\hbar$. The singularity of the Dirac wave function at the origin makes $\Delta p^2$ in finite at $\gamma = 1/2$ (ultra-relativistic regime $d = \infty$). This value of $\gamma$ gives $Z = 116$ which corresponds to the very end of the Mendeleev table.

4.2. Dirac hopfion

The localized solution the Dirac equation for a free electron $\Psi_H$, called the Dirac hopfion [19], can be represented as the following Fourier integral ($\hbar = 1$):

$$\Psi_H(\mathbf{r}, t) = N_H \int d^3p \, \hat{\Psi}(\mathbf{p}) e^{-i\mathbf{p}\mathbf{r} + i\mathbf{p}\cdot\mathbf{a}}, \tag{37}$$

where

$$\hat{\Psi}(\mathbf{p}) = e^{-a\mathbf{p}} \begin{bmatrix} 1 & 0 \\ \frac{E_p - p_x}{m} & \frac{m}{p_x + i\gamma_p} \end{bmatrix}. \tag{38}$$

The integral (38) can be evaluated in terms of Macdonald functions [19] but for our purpose the Fourier representation is sufficient. The normalization coefficient can be expressed in terms of the Macdonald function $K_2$,

$$N_H^2 = \int d^3p |\hat{\Psi}(\mathbf{p})|^2 = \frac{K_2(2ma)}{a}. \tag{39}$$

In the present case, we have a specific function at hand so that we cannot avoid the inclusion of the average value of the momentum $\langle \mathbf{p} \rangle$ in the definition (6) of dispersion.

The relevant integrals in momentum space have been evaluated numerically with Mathematica [20] and the results are shown in figure 2. The nonrelativistic limit $3/2\hbar$ is approached when $a \to \infty$, i.e. when the extension of the wave packet is much larger than the electron Compton wave length.

5. Conclusions

We believe that we have shown that the differences between the relativistic and the nonrelativistic uncertainty relations are substantial. One may trace these differences to the presence in relativistic quantum mechanics of the scale factor, the Compton wave length $\hbar/mc$. This fact is reflected in the appearance of the dimensionless
parameter $d$ in our calculations. The parameter $d$ allows one to compare directly the values of the two dispersions $\Delta r^2$ and $\Delta p^2$. One may tell that we are in the nonrelativistic regime, when the wave function is more squeezed in the momentum space than in the position space. The nonrelativistic limit is reached in two distinct physical situations, when $\Delta r^2 \rightarrow \infty$ or $\Delta p^2 \rightarrow 0$. The dependence of the lower bound on $d$ in the uncertainty relation, as illustrated in figure 1 stands in contrast to the standard nonrelativistic quantum mechanics where all Gaussians saturate the uncertainty relation.

**ORCID iDs**

Iwo Bialynicki-Birula  [https://orcid.org/0000-0002-5195-4133](https://orcid.org/0000-0002-5195-4133)

**References**