Uncertainty relation for focal spots in light beams

Iwo Bialynicki-Birula
Center for Theoretical Physics, Polish Academy of Sciences
Aleja Lotników 32/46, 02-668 Warsaw, Poland

Zofia Bialynicka-Birula
Institute of Physics, Polish Academy of Sciences
Aleja Lotników 32/46, 02-668 Warsaw, Poland

Uncertainty relations for light pulses found in [Phys. Rev. A 86, 022118 (2012)] were derived in the three-dimensional case which emphasized the localization in a volume. Here we derive the uncertainty relation for light beams in the two-dimensional plane perpendicular to the direction of the beam propagation which is more interesting for realistic beams. This uncertainty relation connects the area of the focal spot with the spectrum of transverse photon momenta. The shape of the beam that saturates the uncertainty relation is very close to a Gaussian. The directions of the electric and magnetic field vectors are those of the circularly polarized plane wave. Our uncertainty relation for the focal spot is quite general but we were able to determine the value of the lower bound only for beams made of many photons.

I. INTRODUCTION

Quantum-mechanical properties of photons impose a relationship between the spatial and the spectral properties of light beams. It may be given the form of uncertainty relations derived in \[1, 2\]. The uncertainty relations for light beams follow from the fundamental properties of relativistic quantum systems and they cannot be overcome by technical improvements introduced in the process of the beam formation.

In this paper we analyze the restrictions imposed on the area of the focal spot (beam waist) by the Heisenberg-type uncertainty relation that involves the variance of the transverse wave vector. We fully confirm our intuitive feeling that the better the beam is focused the less certain is the transverse beam momentum. To make the calculations feasible we assume that the beam is intense, i.e. the average number of photons \(N\) is very large. This enables us to simplify the calculations by keeping only the leading terms in the expansion in \(1/N\).

There is some overlap between this work and our previous papers. However, there we mostly dealt with the uncertainty relations for single photons. The treatment of light beams in the last section of Ref. \[2\] was rather sketchy and it dealt only with light pulses and not with light beams since we studied exclusively the three-dimensional case. Here we derive the two-dimensional uncertainty relation which is more appropriate for light beams.

There are many different ways to express the physical uncertainty relations in terms of mathematical inequalities. For example, in the nonrelativistic quantum mechanics we have the standard Heisenberg relations but also the entropic uncertainty relations. In the case of beams of light we have an almost universal measure of the directional spread of wave vectors but many measures of the spatial spread. Siegman \[3\] listed seven possible definitions of the beam width “that have been suggested or used for optical beams in the past.” However, only one or two might be useful for the derivation of uncertainty relations. In this context we must mention the uncertainty relations derived earlier by M. A. Alonso and his collaborators \[4–7\] that are based on a completely different measure of the spatial spread. We shall compare their results with ours in the Conclusions.

To introduce the main concepts, let us consider a well collimated beam of light centered on the \(z\) axis. Following \[2\], we shall measure the beam width with the use of the center of energy operator \(\hat{R}\) introduced almost eighty years ago by Born and Infeld \[8\]. Even though the components of this operator do not commute, it is still the best available substitute for the nonexistent genuine photon position operator. The size of the beam waist will be measured by the variance of the center of energy operator \(\hat{R}\) in the transverse plane,

\[
\Delta R^2 = \langle R^2 \rangle - \langle R \rangle^2
\]

and the spread of the transverse momentum by the variance of the transverse momentum operator:

\[
\Delta P^2 = \langle P^2 \rangle - \langle P \rangle^2.
\]

We prove that the following uncertainty relation holds:

\[
\sqrt{\Delta R^2} \sqrt{\Delta P^2} \geq \gamma \hbar.
\]

The dimensionless parameter \(\gamma\) depends on the mean value \(\langle k_z \rangle\) of the component of the wave vector in the beam direction. In the limit, when \(\langle k_z \rangle\) tends to infinity, \(\gamma\) tends to 1. In the process of proving this relation we find the wave fields that saturate the inequality.

The use of the center of energy operator to characterize the spatial extension of a single photon may seem awkward. However, its use for intense light beams should
not raise doubts because it is the energy distribution in a beam that gives the best characteristic of the beam shape.

II. CENTER OF THE ENERGY OPERATOR

In classical electrodynamics we define the center of the energy $R$ of the electromagnetic field as the following ratio:

$$R = \frac{\int d^3r \, r \mathcal{E}(r)}{\int d^3r \, \mathcal{E}(r)},$$

where $\mathcal{E}(r)$ is the energy density. In quantum theory the energy density becomes an operator and to obtain a Hermitian operator $\hat{R}$ we have to symmetrize the product of two noncommuting operators,

$$\hat{R} = \frac{1}{2\hbar} \hat{N} + \frac{1}{2\hbar},$$

where we identified the total energy operator with the Hamiltonian,

$$\hat{H} = \int d^3r \, \hat{\mathcal{E}}(r),$$

and we denoted the first moment of the energy by $\hat{N}$,

$$\hat{N} = \int d^3r \, r \hat{\mathcal{E}}(r).$$

Since the momentum operator of the electromagnetic field acts as the generator of translations,

$$\left[ \hat{P}, \hat{\mathcal{E}}(r) \right] = i\hbar \nabla \hat{\mathcal{E}}(r),$$

the operators $\hat{R}$ and $\hat{P}$ obey the canonical commutation relations for position and momentum,

$$[\hat{R}_i, \hat{P}_j] = i\hbar \delta_{ij}.$$  \hspace{1cm} (9)

There is, however, one disconcerting property of the center of the energy operator: its components do not commute. This follows directly from the commutator of the energy densities,

$$[\hat{\mathcal{E}}(r), \hat{\mathcal{E}}(r')] = -i\hbar c^2 \left[ \hat{\mathcal{P}}^k(r) - \hat{\mathcal{P}}^k(r') \right] \partial_k \delta(r - r'),$$

where $\hat{\mathcal{P}}(r)$ is the density of momentum. This relation was called by Schwinger \[8\] “the most fundamental equation of relativistic quantum field theory” since it must hold for every relativistic quantum system. With its use we obtain first

$$\hat{N} \times \hat{N} = -i\hbar c^2 \hat{M},$$

and then \[8\],

$$\hat{R} \times \hat{R} = -i\hbar c^2 \frac{1}{\hbar} \hat{S} \frac{1}{\hbar} = \frac{1}{\hbar^2} \hat{S},$$

where the spin operator $\hat{S}$ is defined as the difference between the total angular momentum $\hat{M}$,

$$\hat{M} = \int d^3r \, \hat{r} \times \hat{\mathcal{P}}(r),$$

and the orbital angular momentum $\hat{R} \times \hat{P}$. In the case of photons (or the electromagnetic field) the role of spin is played by helicity.

It is worthwhile to note that the complete list of commutators for the operators $\hat{H}, \hat{P}, \hat{M},$ and $\hat{N}$ coincides with the commutator algebra for the Poincaré group with $\hat{N}$ being the generator of Lorentz transformations. Thus, noncommutativity of the components of $\hat{R}$ is a direct consequence of the noncommutativity of Lorentz transformations. This noncommutativity does not preclude the use of $\hat{R}$ to characterize the spatial extension of the system. In a relativistic system there is no other quantity that can better serve as a replacement for the nonrelativistic position.

III. LIGHT BEAMS AS COHERENT STATES OF THE ELECTROMAGNETIC FIELD

To describe light beams we choose the coherent states: the most classical states of the electromagnetic field. Coherent states $\langle \text{coh} \rangle$ are constructed \[10\] by acting on the vacuum state vector with the unitary Glauber displacement operator $D$,

$$\langle \text{coh} \rangle = D|0\rangle,$$  \hspace{1cm} (14)

$$D = \exp \left( \sqrt{\hat{N}} \sum \left[ f^\dagger_{\lambda}(\mathbf{k}) \hat{a}^\lambda_{\lambda}(\mathbf{k}) - f^\dagger_{\lambda}(\mathbf{k}) \hat{a}^\lambda_{-\lambda}(\mathbf{k}) \right] \right),$$  \hspace{1cm} (15)

where the symbol $\sum$ stands for both the integration over all wave vectors $\int d^3k/k$ and the summation over two circular polarizations $\lambda = \pm 1$. The factor $1/k$ transforms the volume element $d^3k$ into the relativistically invariant measure $d^3k/k$ on the light-cone \[11\].

The mode functions $f_{\lambda}(\mathbf{k})$ fully describe the shape of the light beam. We assume the invariant normalization of the mode function,

$$\int \left| f_{\lambda}(\mathbf{k}) \right|^2 = 1.$$  \hspace{1cm} (16)

With this normalization of $f_{\lambda}(\mathbf{k})$ the number $N$ appearing in \[14\] is equal to the expectation value of the photon number operator $\hat{N}$,

$$\hat{N} = \sum \left| a^\dagger_{\lambda}(\mathbf{k}) a_{\lambda}(\mathbf{k}) \right|^2.$$  \hspace{1cm} (17)
in the coherent state, \( N = \langle \hat{N} \rangle \). Our relativistic convention implies that the commutation relations between the annihilation and creation operators contain an extra factor of \( \hbar \),

\[
\left[ a_N(k), a_N^\dagger(k') \right] = \delta_{N,k} \delta^{(3)}(k-k'). \tag{18}
\]

The displacement operator shifts the creation and annihilation operators by the following c-number terms:

\[
D^\dagger a_N^\dagger(k) D = a_N(k) + \sqrt{N} f_N^\dagger(k), \tag{19a}
\]

\[
D^\dagger a_N(k) D = a_N(k) + \sqrt{N} f_N(k). \tag{19b}
\]

The explicit appearance of \( N \) in these formulas simplifies the extraction of the leading terms in the uncertainty relations for intense beams.

IV. MEASURES OF UNCERTAINTY IN POSITION AND MOMENTUM

In order to evaluate the variances (11) and (2) we must express \( \hat{H}, \hat{P}_\perp, \) and \( \hat{N}_\perp \) in terms of creation and annihilation operators [12, 13],

\[
\hat{H} = \sum \hbar \omega a_k^\dagger(k)a_k(k), \tag{20a}
\]

\[
\hat{P}_\perp = \sum \hbar k a_k^\dagger(k)a_k(k), \tag{20b}
\]

\[
\hat{N}_\perp = \sum \hbar a_k^\dagger(k)iD_{\lambda}a_k(k), \tag{20c}
\]

where \( D_\lambda \) is the covariant derivative on the light cone.

\[
D_\lambda = \nabla_k - i\lambda \alpha(k),
\]

and \( \alpha(k) \) is the analog of the affine connection,

\[
\alpha(k) = \frac{k_z(-k_y,k_x,0)}{k(k_x^2 + k_y^2)}. \tag{22}
\]

The expression for \( \Delta P_\perp^2 \) obtained with the use of (19) contains only one term,

\[
\Delta P_\perp^2 = N \sum \hbar^2 k_\perp^2 |f_\lambda(k)|^2, \tag{23}
\]

because the terms proportional to \( N^2 \) cancel between the two parts of the variance present in (2).

In order to calculate \( \Delta R_\perp^2 \) we need the following formulas for \( D^\dagger \hat{N}_\perp D \) and \( D^\dagger (1/\hat{H}) D \):

\[
D^\dagger \hat{N}_\perp D = N \left[ \sum \hbar \omega f_\lambda(k)iD_{\lambda\perp}f_\lambda(k) + \frac{1}{\sqrt{N}} \sum \hbar \omega \left( a_k^\dagger(k)iD_{\lambda\perp}f_\lambda(k) + f_\lambda^\dagger(k)iD_{\lambda\perp}a_k(k) \right) + \mathcal{O}\left( \frac{1}{N} \right) \right], \tag{24a}
\]

\[
D^\dagger \frac{h\hat{c}}{\hat{H}} D = \frac{1}{N\hbar} \left[ 1 - \frac{1}{N\hbar} \sum k \left( a_k^\dagger(k)f_\lambda(k) + f_\lambda^\dagger(k)a_k(k) \right) + \mathcal{O}\left( \frac{1}{N} \right) \right], \tag{24b}
\]

where \( \mathcal{H} \) is the average value of \( k \) per one photon in the coherent state,

\[
\mathcal{H} = \sum k |f_\lambda(k)|^2. \tag{25}
\]

The evaluation of \( \Delta R_\perp^2 \) in the lowest order is now simple but tedious. Again the leading terms cancel between the two terms in the variance (11) resulting in:

\[
\Delta R_\perp^2 = \frac{1}{N\mathcal{H}^2} \sum k^2 \left[ (iD_{\lambda\perp} - \mathcal{R}_\perp)f_\lambda(k) \right]^* (iD_{\lambda\perp} - \mathcal{R}_\perp)f_\lambda(k), \tag{26}
\]

where

\[
\mathcal{R}_\perp = \frac{1}{\mathcal{H}} \sum k f_\perp^\dagger(k)iD_{\lambda\perp}f_\perp(k). \tag{27}
\]

We can eliminate \( \mathcal{R}_\perp \) by changing the phase of \( f_\lambda(k) \).

\[
f_\lambda(k) \rightarrow e^{-i\mathcal{R}_\perp} f_\lambda(k). \tag{28}
\]

This change of the phase in momentum space is equivalent to the translation by \( \mathcal{R}_\perp \) in the coordinate space. Since \( \langle \mathcal{R}_\perp \rangle = \mathcal{R}_\perp + \mathcal{O}\left( \frac{1}{N} \right) \), after the translation the origin of the coordinate system coincides with the center of energy. Alternatively, we could have chosen the coordinate system in this way from the very beginning and then \( \langle \mathcal{R}_\perp \rangle \) would be equal to 0.

Finally, the left hand side of the uncertainty relation takes on the form:

\[
\gamma_\perp^2 = \frac{\Delta R_\perp^2 \Delta P_\perp^2}{\hbar^2} = \frac{\text{certain expression}}{\text{certain expression}}, \tag{29}
\]

We replaced momenta by the corresponding wave vectors since for intense light beams the Planck constant does not play a significant role. In the next section we apply the variational procedure to find the minimum of \( \gamma_\perp \).
V. THE VARIATIONAL EQUATION

Up to this point we followed the general procedure developed previously for the three-dimensional case. However, the formula (29) for $\gamma^2$ differs from the formula (66) for $\gamma^2$ in [2] so that the variational equation for $f_\lambda(k)$ will now be different. Variation of $\gamma^2_\perp$ with respect to $f_\lambda^2(k)$ leads to an equation that allows for the separation of variables in polar coordinates. After the substitution $f_\lambda(k) = \exp(iM\phi)f_\lambda(k_\perp,k_z)$ we obtain the following equation for $f_\lambda(k_\perp,k_z)$:

$$Df_\lambda(k_\perp,k_z) = 0,$$

where $D$ is the following differential operator:

$$D = -\partial^2_{\kappa_\perp} - \left(\frac{1}{\kappa_\perp} + \frac{\kappa_\perp}{\kappa^2_\perp + \kappa^2_z}\right)\partial_{\kappa_\perp} + \frac{1}{\kappa^2_\perp} \left(\frac{\kappa^3_z}{\sqrt{\kappa^2_\perp + \kappa^2_z}} - \lambda M\right)^2 + \gamma^2_\perp \left(\frac{\kappa^2_z}{\kappa^2_\perp + \kappa^2_z} - \frac{2}{\sqrt{\kappa^2_\perp + \kappa^2_z}}\right).$$

We introduced here the dimensionless variables,

$$\kappa_\perp = \frac{k_\perp \mathcal{H}}{\Delta k^2_\perp}, \quad \kappa_z = \frac{k_z \mathcal{H}}{\Delta k^2_\perp}.$$  \hspace{1cm} (32)

It is sufficient to consider one polarization at a time, say $\lambda = 1$, because the two functions $f_\pm$ satisfy separate equations. The variational equation (30) contains $\kappa_z$ only as a parameter. This equation is linear, therefore, every solution can be multiplied by an arbitrary function of $\kappa_z$, say $A(\kappa_z)$. Such a multiplicative factor is needed, in particular, to secure normalizability.

VI. SLIM BEAMS

We shall characterize beam shapes with the use of the following dimensional parameters:

- The standard deviation $d_\perp = \sqrt{\Delta k^2_\perp}$ that measures the spread of the transverse components of the wave vector (the beam width in momentum space).
- The average value $\langle \kappa_z \rangle$ of the longitudinal component of the wave vector.
- The standard deviation $d_z = \sqrt{\Delta k^2_z}$ that measures the width of the distribution of $k_z$, as determined mainly by the amplitude $A(\kappa_z)$.

Two dimensionless ratios: $s = d_\perp/\langle \kappa_z \rangle$ and $w = d_z/\langle \kappa_z \rangle$ characterize the properties of the beam in a scale-independent way. The parameter $s$ determines how slim the beam is. The parameter $w$ determines how well the average value $\langle \kappa_z \rangle$ represents all the values of $k_z$.

In most cases one is interested in (almost) monochromatic beams. Therefore, both $s$ and $w$ must be very small as compared to 1. We shall call the beams that satisfy the conditions $s \ll 1$ and $w \ll 1$ the slim beams. Note that for slim beams $\langle \kappa_z \rangle$ behaves as $1/s^2$ because $\mathcal{H}$ in (32) is roughly equal to $\langle \kappa_z \rangle$. In what follows we often consider the limit when $w \to 0$, i.e. we neglect the width of the distribution $d_z$. We may then identify $\kappa_z$ in the variational equation with its average value $\langle \kappa_z \rangle$, dropping the average symbol $\langle \rangle$ altogether. Sometimes, however, for the sake of clarity we will keep the distinction.

It is worth noting that the $1/\kappa^2_z$ term in the variational equation (30) corresponds to a centrifugal repulsion that pushes the function $f(\kappa^2_\perp,\kappa_z)$ away from the region of small $1/\kappa^2_z$. The presence of this term clearly increases the variance $\Delta k^2_z$, resulting in the increase of the lower bound in the uncertainty relation. To decrease the role of this repulsion we should take $M = \lambda = 1$ or $M = \lambda = -1$ and a large value of $\langle \kappa_z \rangle$. The other choice $M = 0$ and $\kappa_z = 0$ also eliminates the centrifugal term but the vanishing of $\kappa_z$ means that we no longer deal with a beam.

In Fig. 1 we show the lower bound $\gamma_\perp$ as a function of the average value of longitudinal component of the wave vector. It is seen that the value $\gamma_\perp = 1$ (characteristic for nonrelativistic quantum mechanics in two dimensions) is obtained when $\kappa_z \to \infty$. For $\kappa_z = 0$ the variational equation (31) can be solved analytically giving $\gamma_\perp(0) = (1 + \sqrt{5})/2 \approx 1.618$. In Fig. 2 we show for $\kappa_z = 100$ the function $f(\kappa_\perp,\kappa_z)$ that solves the variational equation (with the normalization such
that \( f(0, \kappa_z) = 1 \) compared with the Gaussian function \( g(\kappa_\perp, \kappa_z) \),

\[
g(\kappa_\perp, \kappa_z) = \exp\left(-\frac{\kappa_\perp^2}{2\kappa_z}\right) = \exp\left(-\frac{k_\perp^2}{2\Delta k_\perp^2 \kappa_z}\right). \tag{33}
\]

These two functions are practically identical. The Gaussian function gives rise to the numerical solution \( f(\kappa_\perp) \) of the variational equation corresponding to the value of the dimensionless parameter \( (\kappa_z) = 100 \). The difference between these two curves is hidden in the line width.

The expectation value of the field operator in the coherent state \( |\alpha\rangle \) gives the connection between the field and the functions \( f_\perp(k) \),

\[
F(r, t) = \frac{D}{\sqrt{2\epsilon}} + i \frac{B}{\sqrt{2\mu}} = \sqrt{\sum} \frac{d^3k}{(2\pi)^{3/2}} e(k) \times \left[ a_+^\dagger(k) e^{i k \cdot r - i\omega t} + a_-^\dagger(k) e^{-i k \cdot r + i\omega t} \right], \tag{34}
\]

where

\[
e(k) = \frac{1}{\sqrt{2k_\perp k_\perp^2 + k_y^2}} \left[ \begin{array}{c} -k_y k_z + i k k_y \\ -k_y k_z - i k k_x \\ k_\perp^2 + k_z^2 \end{array} \right]. \tag{35}
\]

The expectation value of the field operator in the coherent state \( |\alpha\rangle \) gives the connection between the field and the functions \( f_\perp(k) \),

\[
F(r, t) = \frac{D}{\sqrt{2\epsilon}} + i \frac{B}{\sqrt{2\mu}} = \sqrt{\sum} \frac{d^3k}{(2\pi)^{3/2}} e(k) \times \left[ f_\perp^+(k) e^{i k \cdot r - i\omega t} + f_\perp^*(k) e^{-i k \cdot r + i\omega t} \right], \tag{36}
\]

We will take only the part with \( \lambda = 1 \). Since the Gaussian function serves as a very good approximation, as seen in Fig. 2 we shall use this analytic representation to determine the field in the coherent state that minimizes the uncertainty relation. We will again consider slim beams not only to make the Gaussian approximation applicable but also to make the integration in \( (36) \) feasible. First, we use the smallness of \( w \) to perform the integration over \( k_\perp \) and substitute the average value \( \langle k_\perp \rangle \) in place of the integration variable \( k_\perp \). In the remaining two-dimensional integral we choose the polar coordinates. Since we have chosen \( \lambda = 1 \) we must also take \( m = 1 \) and we are left with the function \( e^{i\varphi} e^{-\alpha k^2_\perp / 2} \). The parameter \( \alpha \) (as given by the formula \( (33) \)) for slim beams becomes \( \alpha = 1/\Delta k_\perp^2 \). Thus, we end up with the following integral over the polar variables \( \phi \) and \( k_\perp \):

\[
F(\rho, \varphi, z, t) = \int_0^\infty dk_\perp k_\perp \int_0^{2\pi} d\phi \left[ \frac{k_\perp \cos \phi - i k \sin \phi}{k_\perp \sin \phi + i k \cos \phi} \right] \times e^{i\varphi} e^{-\alpha k^2_\perp / 2} e^{i\rho \cos(\phi - \varphi)} e^{i(k_\perp - z)e^{-i\omega t}}, \tag{37}
\]

where \( \phi - \varphi \) is the angle between the vectors \( k_\perp \) and \( r_\perp \), \( k = \sqrt{k^2_\perp + \langle k^2_z \rangle} \). We dropped an irrelevant overall field amplitude. After the shift of the integration variable \( \phi \to \phi + \varphi \) the integration over \( \phi \) can be performed (Eq. 8.411 of [15]), leading to the following integral over \( k_\perp \) involving Bessel functions:

\[
F(\rho, \varphi, z, t) = \int_0^\infty dk_\perp k_\perp \left[ e^{i(k_\perp - z)e^{-i\omega t}} \right] \times e^{-\alpha k^2_\perp / 2}, \tag{38}
\]

where \( k_\perp = (k \pm \langle k_\perp \rangle) / 2 \). Note that the integrand in this formula is nothing else but the Bessel beam for \( M = 1 \), cf., for example [14]. Therefore the field is the superposition of Bessel beams with different \( k_\perp \) weighted by the Gaussian function.

The properties of slim beams enable us to replace \( k \) by \( \langle k_\perp \rangle \) and to disregard the terms proportional to \( k_\perp \approx \langle k_\perp \rangle \langle k_\perp \rangle \) as compared to the terms proportional to \( k_\perp \approx 2\langle k_\perp \rangle \). The integration over \( k_\perp \) can now be done (Eq. 6.631 of [15]) and the final result is (up to normal-
The plane is axially symmetric and has the Gaussian form, diffractive. The energy distribution in the transverse plane is radially polarized beams do not have the smallest possible focal area. Therefore, radially polarized beams do not have the smallest possible focal area.

Thus, within our approximation, this beam is non-diffractive. The energy distribution in the transverse plane is axially symmetric and has the Gaussian form, so called radially polarized beams studied in [17]. There-fore, radially polarized beams do not have the smallest possible focal area.

\[ F(\rho, \varphi, z, t) = \left[ \frac{1}{i \omega z^2} \right] e^{-\rho^2 \Delta k_z^2 / 2} e^{i(k_z)(z-ct)}. \]  

(39)

Thus, within our approximation, this beam is non-diffractive. The energy distribution in the transverse plane is radially polarized beams do not have the smallest possible focal area.

\[ E(\rho) = F^* \cdot F = \left[ 2 + \left( \frac{\rho \Delta k_z^2}{\langle k_z \rangle} \right)^2 \right] e^{-\rho^2 \Delta k_z^2}. \]  

(40)

On the beam axis the electric field and the magnetic field are those of the circularly polarized plane wave. As we move away from the center the longitudinal component appears (it is needed to make the field divergence-less). However, the longitudinal component is practically negligible as compared with the transverse components because its maximal value \( \sqrt{\Delta k_z^2} / 2 \langle k_z \rangle \) is much smaller than 1. The electric and magnetic field vectors at each point form an orthogonal pair that has the same orientation in the whole transverse plane, as shown in Fig. 3. The field configuration is markedly different from the so called radially polarized beams studied in [17]. Therefore, radially polarized beams do not have the smallest possible focal area.

**VIII. CONCLUSIONS**

We derived the uncertainty relation that imposes the lower limit on the focusing area. The size of this area is determined by the distribution of energy (beam intensity). The lowest value of the focal area is determined by the spectral properties of the beam; it is inversely proportional to the variance of the transverse components of the wave vectors present in the beam. In the limit of very slim beams the uncertainty relation becomes the same as the two-dimensional Heisenberg uncertainty relation in nonrelativistic quantum mechanics. We also determined the shape of the beam that saturates the uncertainty relation, i.e. the beam that is maximally focused given the value of the variance of transverse wave vectors. The energy density of this beam has a Gaussian shape while the directions of the field vectors (except for a small longitudinal component) are the same as in the circularly polarized plane wave (Fig. 3).

The main difference between our results and those described in Refs. [3–7] is the use of a different measure of the spatial spread. The spread is measured in these works in a novel way (it is not on the Siegman’s list), namely, by using the square of the angular part of angular momentum. This measure can only be applied to monochromatic beams of light (cf. Eq. (8) of Ref. [3]). In contrast, we use the measure of the beam width which is not restricted to the ideal case of monochromatic beams. Since it is based on the variance of the position operator (defined as the center of energy) our measure is as close as possible to the standard quantum mechanical definition. It seems to be also directly related to the observed intensity distribution in the beam. The second difference is the use of the covariant derivative on the light cone. The use of plain derivatives leads to the violation of divergence condition as seen in the formula (12) of Ref. [7]. Also, our uncertainty relation seems to be more fundamental since it is rooted in the relativistic properties of the electromagnetic field but that does not necessarily mean that it would be more useful in practical applications.

**ACKNOWLEDGMENTS**

This research was partly supported by the grant from the Polish Ministry of Science and Higher Education for the years 2013–2016.

**Appendix A**

As a natural measure of the difference between the Gaussian function and the exact solution of the variational equation we take the ratio \( I(k_z) \) of the integral of the square of the function \( Dg(k_{\perp}, k_z) \) divided by the norm of the Gaussian function,

\[ I(k_z) = \frac{\int_D ds_{\perp} [Dg(k_{\perp}, k_z)]^2}{\int_D ds_{\perp} [g(k_{\perp}, k_z)]^2}. \]  

(A1)

Obviously, when this integral is equal to 0, the func-
dimensionless parameter $\kappa_z$ is very well approximated by the asymptotic formula \[ (A4) \]
which can be evaluated in a closed form \[ (A2) \],
\[ I(\kappa_z) = c_0 + c_1 e^{\kappa_z} K_0(\kappa_z/2) + c_2 e^{\kappa_z/2} K_1(\kappa_z/2) + c_3 e^{\kappa_z} \text{erfc}(\kappa_z), \]
where $K_n$ is the Macdonald function, erfc is the complementary error function, and
\[ c_0 = \frac{204 - 175\kappa_z + 120\kappa_z^2 + 60\kappa_z^3 + 12\kappa_z^4}{12\kappa_z^2}, \]
\[ c_1 = \frac{8 - 18\kappa_z - 12\kappa_z^2}{3\kappa_z^{3/2} \sqrt{\pi}}, \]
\[ c_2 = \frac{-40 + 32\kappa_z + 6\kappa_z^2 + 12\kappa_z^3}{3\kappa_z^{5/2} \sqrt{\pi}}, \]
\[ c_3 = \sqrt{\pi} \frac{17 - 20\kappa_z - 5\kappa_z^2 - 24\kappa_z^3 - 11\kappa_z^4 - 2\kappa_z^5}{2\kappa_z^{7/2}}. \]
The results are shown in Fig. 4 where we plotted the exact function $I(\kappa_z)$ and its asymptotic form:
\[ I(\kappa_z) \approx \frac{33}{16\kappa_z^4}, \]
{
[1] I. Białynicki-Birula and Z. Białynicka-Birula, Phys. Rev. Lett. 108, 140401 (2012).
[2] I. Białynicki-Birula and Z. Białynicka-Birula, Phys. Rev. A 86, 022118 (2012).
[3] A. E. Siegman, in DPSS (Diode Pumped Solid State) Lasers: Applications and Issues, M. Dowley, ed., Vol. 17 of OSA Trends in Optics and Photonics (Optical Society of America, 1998), paper MQ1.
[4] M. A. Alonso and G. W. Forbes, JOSA A 17, 2391 (2000).
[5] M. A. Alonso, R. Borghi and M. Santarsiero, JOSA A 23, 691 (2006).
[6] M. A. Alonso, R. Borghi and M. Santarsiero, JOSA A 23, 701 (2006).
[7] M. A. Alonso J. Opt. 13, 064016 (2011).
[8] M. Born and L. Infeld, Proc. Roy. Soc. Lond. A 150, 141 (1935).
[9] J. Schwinger, Phys. Rev. 130, 406 (1963).
[10] R. J. Glauber, Phys. Rev. 131, 2766 (1963); U. M. Titulaer and R. J. Glauber, Phys. Rev. 145, 1041 (1966); Brian J. Smith and M. G. Raymer, New J. Phys. 9, 414 (2007).
[11] In our previous paper [2] the symbol $\mathcal{Y}$ did not include the factor $1/k$.
[12] I. Białynicki-Birula and Z. Białynicka-Birula, Quantum Electrodynamics (Pergamon, Oxford, 1975), Chap. 9; I. Białynicki-Birula and Z. Białynicka-Birula, Phys. Rev. D 35, 2383 (1987); I. Białynicki-Birula and Z. Białynicka-Birula, J. Opt. 13, 064014 (2011).
[13] I. Białynicki-Birula and Z. Białynicka-Birula, J. Phys. A 46, 053001 (2013).
[14] I. Białynicki-Birula and Z. Białynicka-Birula, Opt. Comm. 264, 342 (2006).
[15] I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Series, and Products, (Academic Press, New York, 2000).
[16] Wolfram Research, Inc., Mathematica, Version 9, Champaign, IL (2012).
[17] S. Quabis, R. Dorn, M. Eberler, O. Glockl, and G. Leuchs, Opt. Commun. 179, 1 (2000); K. S. Youngworth and T. G. Brown, Opt. Express 7, 77 (2000); R. Dorn, S. Quabis, and G. Leuchs, Phys. Rev. Lett. 91, 233901 (2003); S. Yan and B. Yao, Phys. Rev. A 77, 023827 (2008); S. Orlov and U. Peschel, Phys. Rev. A 82, 063820 (2010).]