Minimum uncertainty solutions for partially coherent fields and quantum mixed states

To cite this article: M Alonso et al 2014 New J. Phys. 16 123023

View the article online for updates and enhancements.
Minimum uncertainty solutions for partially coherent fields and quantum mixed states

M Alonso\(^1,2\), T Setälä\(^3\) and A T Friberg\(^3\)

\(^1\)The Institute of Optics, University of Rochester, Rochester NY 14627, USA
\(^2\)Center for Coherence and Quantum Optics, University of Rochester, Rochester NY 14627, USA
\(^3\)Institute of Photonics, University of Eastern Finland, PO Box 111, FI-80101 Joensuu, Finland
E-mail: miguel.alonso@rochester.edu

Received 6 October 2014
Accepted for publication 14 November 2014
Published 8 December 2014

Abstract
A general prescription is given for finding uncertainty relations that dictate the lower bounds on the measures of spread corresponding to two different representations of a partially coherent wave field or mixed quantum state, for a given measure of overall coherence or purity. In particular it is shown that the coherent modes of the fields/states that achieve the lower bounds are independent of the measure of purity being used, and that this measure determines only the amount in which these modes contribute. Our results are important in the design of optical systems with partially coherent light and in quantum mixed states, for which maximal joint localization is desired. These ideas are illustrated for the case of optical beams, pulses propagating in dispersive media and quantum phase.

Keywords: uncertainty relations, coherence, quantum states

1. Introduction

Uncertainty relations provide fundamental lower bounds to the measures of spread that characterize a physical state in different representations [1–3]. The best-known example is the Heisenberg uncertainty principle [4], which states that the product of the rms spreads in position

\[
\Delta x \Delta p \geq \frac{\hbar}{2}
\]

where \(\Delta x\) and \(\Delta p\) are the standard deviations of the position and momentum measurements, and \(\hbar\) is the reduced Planck constant. These relations are essential in quantum mechanics, and their generalizations to mixed states and partially coherent fields are of great importance in optics and quantum information science.
and momentum of a quantum particle has a lower bound. Relations of this type, however, are
not limited to the quantum world; similar relations exist for classical systems. For example, the
Fourier space–bandwidth product inequality provides a lower bound for the temporal duration
and frequency spread of time signals, like sounds or light pulses [1]. In the spatial domain,
analogous relations exist between the waist width and the directional spread of a paraxial optical
or acoustic beam [5]. In all these examples, the two spreads characterizing the system
correspond to the standard deviations of Fourier-conjugate variables (position/momentum, time/
frequency, transverse position/direction).

Uncertainty-type relations can be derived also for situations where the two representations
of interest are not related through continuous Fourier transformation. For example, the relation
between the phase and the number of photons in a quantum-optical mode is given by a (one-
sided) Fourier series relation, and several alternative forms of the uncertainty relation for this
situation have been proposed [6–18]. Analogous relations exist between the angular distribution
of a beam at a transverse plane and its orbital angular momentum [19–22], or in the case of the
spatial and directional spreads of wave fields beyond the paraxial approximation [23–30].
Finally, the relation between the initial temporal spread of an optical pulse traveling through a
dispersive medium and the asymptotic rate of increase of the temporal spread over long
propagation distances are related through the standard Heisenberg relation only when the
dispersion of the medium is linear [31], and a more complicated relation exists for general
lossless dispersion [32].

The relations described earlier can be derived in at least two ways: through Cauchy–
Schwarz inequalities and commutation rules [1–4, 7, 13, 15, 16, 18, 24], and through variational
methods [6, 10, 14, 17, 26, 27, 30]. The first approach is perhaps more elegant, given its direct
connection to the algebraic properties of the system. However, it is able to give the exact lower
bounds (and the systems that achieve them) only for very limited situations, including the
standard case where the two representations of interest are related by Fourier transformation.
In situations where the commutator of the two quantities of interest is not just a number, this
approach becomes more complicated. The variational approach, on the other hand, always gives
the true lower bounds and the states that achieve them, although it often requires the numerical
solution of a Sturm–Liouville differential problem. In this approach, the above-mentioned
commutator is not used explicitly (although of course the resulting minimum-uncertainty states
do reflect the type of commutation of the two operators).

In quantum mechanics, one is often interested in mixed states, i.e., states where the density
matrix cannot be factorized as the product of a wavefunction and its complex conjugate [33].
Analogously, in classical optics, it is often convenient to use a statistical description of the
optical field that accounts for random fluctuations. Fields of this type are referred to as partially
coherent and are described, within the theory of second-order coherence, in terms of a cross-
correlation that depends on two spatial (and/or temporal) arguments [34–36]. This cross-
correlation cannot be factored as the outer product of a field and its complex conjugate except in
the case of purely coherent fields. Measures for the overall level of coherence or purity of
stochastic fields and mixed wavefunctions have been proposed [33, 37, 38], and their effect on
the uncertainty relations for paraxial fields has been studied in the scalar [38–42] and
electromagnetic [43] cases.

In this work, we give a general prescription for finding the lower bounds of any two
measures of spread, corresponding to two representations of the system, for the case of partially
coherent fields or mixed states for a given value of some definition of overall measure of
coherence/purity. The first important result we find is that the coherent modes that diagonalize the optimal solutions correspond to the eigenfunctions of the Sturm–Liouville equation used in the variational approach for finding the optimal coherent fields. Therefore, these coherent modes of the optimal fields are independent of which specific definition of the overall measure of coherence is being used. The second important result is that the amount in which each of these modes contributes to the optimal states depends only on the overall measure of coherence/purity being used, and not on the definitions of spread being used. We illustrate this procedure with several examples.

2. Minimization of the spreads for coherent fields and pure states

Consider first the case of coherent fields or pure states. Let these states depend on a variable \( p \) (which can be discrete, continuous, or a multidimensional mix of the two), so they are described by a complex function \( A(p) \). This function might correspond to the field itself in physical space or to a unitary transformation (e.g., the Fourier transform) of it. Let us denote by \( \Delta_a \) a generic measure of spread (or ‘uncertainty’) of the field or state in the variable associated with a given representation. This measure is defined according to

\[
\int \Delta \Phi = \mathbb{E}_a \left[ \hat{\mathcal{O}}_a A^*(p) A(p) \right],
\]

where \( \mathbb{E}_a \) is some monotonically-growing function and \( \hat{\mathcal{O}}_a \) is an operator related to the variable in question, the asterisk denotes complex conjugation, and \( \Phi \) is the square norm of the function, given by

\[
\Phi = \int A^*(p) A(p) dp.
\]

(For example, in the case of the standard second moment of an operator \( \hat{a} \), \( \hat{\mathcal{O}}_a = (\hat{a} - \bar{a})^2 \), where \( \bar{a} \) is the average/expected value of \( \hat{a} \), and \( g_a(\Delta_a) = \Delta_a^2 \).) The integrals in these equations represent summation over the appropriate range and dimensionality for the physical problem in question; they can also represent discrete sums (e.g., a quantum system of discrete states) or a mixture of continuous and discrete variables (e.g., a classical vector-valued field).

In what follows we consider two spreads \( \Delta_a \) and \( \Delta_b \), and the corresponding two operators \( \hat{\mathcal{O}}_a \) and \( \hat{\mathcal{O}}_b \). The function \( A \) that minimizes \( \Delta_a \) for given \( \Delta_b \) (or vice versa) can be found through the following variational calculation: the functional derivative with respect to \( A^* \) of a monotonically increasing combination of these measures is required to vanish, e.g.

\[
\frac{\delta g_a(\Delta_a)}{\delta A^*} + \nu^2 \frac{\delta g_b(\Delta_b)}{\delta A^*} = 0,
\]

where \( \nu^2 \) is a non-negative Lagrange multiplier. Notice that considering variations in \( A^* \) for ‘fixed’ \( A \) is equivalent to (and simpler than) considering independent variations in both the real and imaginary parts of \( A \). The functional derivative of each measure is easily seen to be given by
\[
\frac{\delta g_i(\Delta_i)}{\delta A^*} = \frac{\hat{\mathcal{O}}_i A(p) - g_i(\Delta_i)A(p)}{\Phi},
\] (4)

for \( i = a, b \), and where the second term comes from the derivative of \( \Phi \). The substitution of this result into equation (3) leads to the following eigenvalue equation:
\[
\left( \hat{\mathcal{O}}_a + \nu^2 \hat{\mathcal{O}}_b \right) A(p) = \lambda A(p),
\] (5)

where the eigenvalue is a linear combination of the widths, i.e., \( \lambda = g_a(\Delta_a) + \nu^2 g_b(\Delta_b) \). Equation (5) has a complete set of solutions, \( \tilde{A}_n \), with corresponding eigenvalues
\[
\tilde{\lambda}_n = g_a(\tilde{\Delta}_{a,n}) + \nu^2 g_b(\tilde{\Delta}_{b,n}),
\] (6)

where \( \tilde{\Delta}_{i,n} \), with \( i = a, b \), are the widths for the eigenfunctions \( \tilde{A}_n \). These eigenvalues are ordered such that \( \tilde{\lambda}_n \leq \tilde{\lambda}_{n+1} \). Assuming that the operators are Hermitian, the functions \( \tilde{A}_n \) can be made to form a complete orthonormal set. From equation (6) it can be seen that the widths \( \Delta_i \) assume their minimum values for the ground state function \( \tilde{A}_0 \) corresponding to the smallest eigenvalue \( \tilde{\lambda}_0 \). For the remaining functions \( \tilde{A}_n \) with \( n > 0 \), the widths \( \Delta_i \) are not minimal, but they are stationary under infinitesimal variations of \( A \).

3. Partially coherent fields and mixed states

We now consider the case of partially coherent fields or quantum mixed states. Assume that \( A \) is a random quantity characterized by the correlation function
\[
\mathcal{A}(p_1, p_2) = \left\langle A^*(p_1)A(p_2) \right\rangle,
\] (7)

where the angle brackets denote an appropriate average. For quantum systems, this quantity corresponds to the density matrix in the representation associated with the variable \( p \) (e.g., the standard density matrix \( \rho(\mathbf{r}_1, \mathbf{r}_2) \) in the representation of position \( \mathbf{r} \)). Any function \( A \) can be expanded as
\[
A(p) = \sum_n a_n A_n(p),
\] (8)

where \( A_n \) constitute an arbitrary orthonormal basis. If \( A \) is a random quantity, the coefficients \( a_n \) in the expansion above are also random. The correlation function in equation (7) can then be expanded as
\[
\mathcal{A}(p_1, p_2) = \sum_{n_1} \sum_{n_2} b_{n_1,n_2} A_{n_1}^*(p_1)A_{n_2}(p_2),
\] (9)

where \( b_{n_1,n_2} \) is the correlation of the expansion coefficients:
\[
b_{n_1,n_2} = \left\langle a_{n_1}^* a_{n_2} \right\rangle.
\] (10)

Given the Hermiticity and non-negative definiteness of \( \mathcal{A} \), the coefficients \( b_{n_1,n_2} \) can be thought of as the elements of a non-negative definite matrix \( \mathbf{B} \) (whose size is infinity for problems where \( p \) is a continuous variable). Notice that, for partially coherent fields
A partially coherent field or a quantum mixed state can be characterized by an overall degree of coherence or purity. Several alternative definitions have been proposed. Perhaps the most standard option is what is known as purity in quantum mechanics [33], whose inverse is sometimes referred to as the Schmidt index [37], which is a measure of the effective number of mutually incoherent modes contributing to the state. This type of measure has been considered within the context of classical partially coherent fields [38–43]. One such measure used in [41] is the square root of the purity, referred to here as \( \mu_2 \):

\[
\mu_2 = \frac{1}{\Phi} \left[ \iint |A(p_1, p_2)|^2 \, dp_1 \, dp_2 \right]^{1/2}. \tag{13}
\]

We consider a more general measure \( \mu \) of the form

\[
\mu = \frac{f(B)}{\Phi}, \tag{14}
\]

where \( f(B) \) is some function of the elements of the matrix \( B \). The measure \( \mu_2 \) in equation (13) is then a special case corresponding to

\[
f(B) = f_2(B) = \sqrt{\text{Tr} \left( B^2 \right)}. \tag{15}
\]

4. Minimization of the spreads for fields and states with given overall level of coherence or purity

In this section, we give a procedure for finding the lower bounds on two spread measures \( \Delta_a \) and \( \Delta_b \) for a field or state of given overall level of coherence or purity \( \mu \), as well as the fields or states that achieve this minimum. This can be done by a generalization of the variational procedure given earlier. For this purpose, it is convenient to use the expansion of the cross-correlation in terms of a basis \( A_n \), as discussed earlier. Then, instead of a functional derivative, we consider the derivative with respect to the expansion coefficients \( b_{n_i,n_2} \) of a monotonically increasing combination of \( \Delta_a \), \( \Delta_b \), and \( \mu \):

\[
\Lambda_a^2 \frac{\partial g_a(\Delta_a)}{\partial b_{n_i,n_2}} + \Lambda_b^2 \frac{\partial g_b(\Delta_b)}{\partial b_{n_i,n_2}} + \frac{\partial \mu}{\partial b_{n_i,n_2}} = 0, \tag{16}
\]

where \( \Lambda_i^2 \), with \( i = a, b \), are again Lagrange multipliers. Substituting equations (11) and (12) in (16), we rewrite this constraint as the following system of equations:
\[
\Lambda_a^2 \int A_{n_1}^*(p) \left( \hat{\partial}_a + \frac{\Lambda_b}{\Lambda_a^2} \hat{\partial}_b \right) A_{n_2}(p) \text{d}p + \frac{\partial f(B)}{\partial b_{n_1,n_2}} = \left[ \Lambda_a^2 g_a(\Delta_a) + \Lambda_b^2 g_b(\Delta_b) + \mu \right] \delta_{n_1,n_2}, \tag{17}
\]

where \(\delta_{n_1,n_2}\) is the Kronecker delta. (Note that we used the fact that \(\Phi \partial \delta = \delta_{n_1,n_2}\)). This equation must be satisfied for all \(n_1, n_2\), even though the right-hand side differs from zero only for \(n_1 = n_2\).

The key element in our derivation is the realization that the integral in left-hand side of equation (17) is diagonalized if we choose, without loss of generality, \(A_n = \tilde{A}_n\), i.e., if we use the basis that resulted from the solution of the coherent problem in section 2. Recall that this basis satisfies \((\hat{\partial}_a + \nu^2 \hat{\partial}_b)\tilde{A}_n = \tilde{\lambda}_n \tilde{A}_n\), with \(\tilde{\lambda}_n = g_a(\tilde{\Delta}_a,n) + \nu^2 g_b(\tilde{\Delta}_b,n)\), where as mentioned earlier \(\tilde{\Delta}_a,n\) are the widths of the eigenfunction \(\tilde{A}_a\). Therefore, by using \(A_n = \tilde{A}_n\) and the orthonormality of the basis functions, the first term in equation (17) simplifies to

\[
\Lambda_a^2 \int A_{n_1}^*(p) \left( \hat{\partial}_a + \frac{\Lambda_b}{\Lambda_a^2} \hat{\partial}_b \right) A_{n_2}(p) \text{d}p = \Lambda_a^2 \tilde{\lambda}_{n_2} \delta_{n_1,n_2}
\]

\[
= \left[ \Lambda_a^2 g_a(\tilde{\Delta}_a,n_2) + \Lambda_b^2 g_b(\tilde{\Delta}_b,n_2) \right] \delta_{n_1,n_2}. \tag{18}
\]

Consequently, (17) becomes

\[
\frac{\partial f(B)}{\partial b_{n_1,n_2}} = \left\{ \Lambda_a^2 \left[ g_a(\Delta_a) - g_a(\Delta_{a,n}) \right] + \Lambda_b^2 \left[ g_b(\Delta_b) - g_b(\Delta_{b,n}) \right] + \mu \right\} \delta_{n_1,n_2}. \tag{19}
\]

This equation holds for any valid definition of \(\mu\). It states that, in this representation, the overall degree of coherence depends only on the diagonal elements of \(B\). This can only be true if \(B\) is a diagonal matrix. Therefore, regardless of the definition of the degree of overall coherence being used, the functions \(\tilde{A}_n\) are what are known as the ‘coherent modes’ [35, section 4.7] of \(A(p_1, p_2)\), i.e., the orthonormal basis under which the expansion of this function is diagonal:

\[
A(p_1, p_2) = \sum_n \tilde{b}_n \tilde{A}_n^*(p_1) \tilde{A}_n(p_2), \tag{20}
\]

where, in this basis, \(\tilde{b}_n = b_{n,n} \geq 0\) and \(B = \tilde{B} = \text{Diag}(\tilde{b}_0, \tilde{b}_1, \ldots)\). Equation (19) implies that the coefficients \(\tilde{b}_n\) are determined from

\[
\frac{\partial f(B)}{\partial b_n} = \Lambda_a^2 \left[ g_a(\Delta_a) - g_a(\Delta_{a,n}) \right] + \Lambda_b^2 \left[ g_b(\Delta_b) - g_b(\Delta_{b,n}) \right] + \mu
\]

\[
= \alpha^2 - \beta^2 (\tilde{\lambda}_n - \tilde{\lambda}_0), \tag{21}
\]

where \(\alpha^2 = \Lambda_a^2 [g_a(\Delta_a) - \tilde{\lambda}_0] + \Lambda_b^2 g_b(\Delta_b) + \mu\) and \(\beta^2 = \Lambda_a^2\) are two non-negative constants.

That is, the weight distribution for the coherent modes is determined by the definition of the overall degree of coherence and by the eigenvalues of the Sturm–Liouville equation. Notice from equation (11) that, in this basis

\[
\Phi = \sum_n \tilde{b}_n. \tag{22}
\]
Let us consider some specific measures of overall (also referred to as global, average, or effective) degrees of coherence or purity. Bastiaans considered a family of overall degrees of coherence, \( \mu_q \), corresponding to the function \( f \) given by [42]

\[
f_q(\mathbf{B}) = \left( \sum_n \tilde{b}_n^q \right)^{1/q},
\]
where \( q > 1 \). For \( q = 2 \) this expression reproduces equation (15). In this case, the left-hand side of (21) gives

\[
\frac{\partial f_q(\mathbf{B})}{\partial \tilde{b}_n} = \left[ \frac{\tilde{b}_n}{f_q(\mathbf{B})} \right]^{q-1},
\]
so that

\[
\tilde{b}_n \propto \left[ 1 - \gamma^2(\tilde{\lambda}_n - \tilde{\lambda}_0) \right]^{\frac{1}{q-1}},
\]
where \( \gamma = \beta/\alpha \). For each value of \( \gamma \) one obtains a different value of the chosen overall degree of coherence. Note, however, that \( \tilde{b}_n \geq 0 \) only for modes where \( \tilde{\lambda}_n \leq \gamma^{-2} + \tilde{\lambda}_0 \), so the coherent mode expansion, equation (20), can only include these modes. Therefore, the value of \( \gamma \) determines how many coherent modes contribute to the optimal field. An alternative measure, \( \mu_e \), associated with the negative of Shannon’s entropy, corresponds to the function \( f \) given by [42]

\[
f_e(\mathbf{B}) = -\sum_n \tilde{b}_n \ln \left( \frac{\tilde{b}_n}{\Phi} \right).
\]
The left-hand side of equation (21) then takes the form

\[
\frac{\partial f_e(\mathbf{B})}{\partial \tilde{b}_n} = -\ln \left( \frac{\tilde{b}_n}{\Phi} \right) - 1 + \sum_{n'} \tilde{b}_{n'} \Phi = -\ln \left( \frac{\tilde{b}_n}{\Phi} \right),
\]
so that

\[
\tilde{b}_n \propto \exp \left( -\beta^2 \tilde{\lambda}_n \right).
\]
In this case the parameter \( \beta \) regulates the value of the overall degree of coherence. Notice that there is no cutoff, since \( \tilde{b}_n \geq 0 \) for all \( n \). This entropic measure of coherence is a limiting case of \( \mu_q \) according to [42] \( \mu_e = \lim_{q \to 1} \frac{1}{q-1} \left( \mu_q - 1 \right) \).

5. Examples

5.1. Fourier transforms and paraxial beams

We start with a problem described by the standard one-dimensional (1D) Fourier transform: a paraxial beam propagating in the \( z \) direction, for which the function \( A(p) \) is the (complex) random field amplitude at the plane \( z = 0 \) and \( p \) is the position coordinate \( x \). (Later, the results will be generalized to beams depending on both \( x \) and \( y \).) The aim is to find, for a fixed overall degree of coherence \( \mu \), the correlation function \( A(x_1, x_2) \) of the optimal beam for which the
intensity’s standard deviation width at the waist and its rate of increase are jointly minimized. The operators associated with these measures are \( \hat{\sigma}_a = x^2 \) (the mean is located at \( x = 0 \)) and \( \hat{\sigma}_b = -(1/k^2) d^2/dx^2 \), respectively, where \( k \) is the wave number. For these operators equation (5) becomes

\[
-A''(x) + V(x)A(x) = \xi A(x),
\]

where \( V(x) = (kx/\nu)^2 \) and \( \xi = (k/\nu)^2 \lambda \). Thus the situation is governed by an equation whose form coincides with the time-independent Schrödinger equation with quadratic potential. The eigenfunctions are therefore Hermite–Gaussian functions which are used as the basis functions in the numerical procedure described in appendix A for finding the eigenfunctions \( \tilde{\lambda}_n(x) \) and the eigenvalues \( \tilde{\lambda}_n \) of equation (5) in a general case. In particular, \( \tilde{\lambda}_n = (2n+1)\tilde{\lambda}_0 \) for any \( \nu \), and therefore equation (25) with \( q = 2 \) gives simply \( \tilde{\delta}_n \propto 1 - 2g_n \), where \( g = \gamma^2 \tilde{\lambda}_0 \). The fact that the weights of the coherent modes must be non-negative then implies that we must only consider modes for which \( n \leq N = \lfloor 1/2\nu \rfloor \), where \( \lfloor \cdot \rfloor \) denotes the floor function. By now using equations (22) and (23) in (14) with \( f = f_2 \), the overall degree of coherence \( \mu_2 \) can be found to be

\[
\mu_2 = \sqrt{\frac{3 - 2(3 - g)gN + 4g^2N^2}{3(N + 1)(1 - gN)^2}}.
\]

Furthermore, since for Hermite–Gaussian functions \( \Delta_{\nu,n} = \sqrt{2n + 1}\Delta_{\nu,0} \) and \( \Delta_{\nu}^2 = \sum \tilde{\delta}_n \Delta_{\nu,n}^2/\Phi \) independently of \( \nu \), the uncertainty relation for fields with specified \( \mu_2 \) can be found to be

\[
\Delta_{\nu} \Delta_{\Phi} \geq \frac{\eta}{2k},
\]

where

\[
\eta = \frac{\sum \tilde{\delta}_n (2n + 1)}{\Phi} = \frac{3(N + 1) - 5gN - 4gN^2}{3(1 - gN)}.
\]

Note that the coherent limit corresponds to \( N = 0 \) (since only one coherent mode is included), and equations (30) and (32) give \( \mu_2 = 1 \) and \( \eta = 1 \), so formula (31) leads to the standard uncertainty relation for Fourier-transform pairs. On the other hand, in the low coherence limit \( N \to \infty \), \( \eta \mu_2^2 \to 8/9 \), as shown in [38]. It turns out that \( 8/9 \leq \eta \mu_2^2 \leq 1 \) for any degree of coherence, so the right-hand side of formula (31) can be replaced with \( 4/9 \mu_2^2 \). Figure 1 shows the coefficients \( \tilde{\delta}_n \) for the degrees of coherence \( \mu_2 = 0.2 \) (for which \( N = 33 \)), \( \mu_2 = 0.6 \) (for which \( N = 3 \)), and \( \mu_2 = 1.0 \) (for which \( N = 0 \)). These coefficients are normalized so that \( \sum \tilde{\delta}_n = \Phi = 1 \).

Let us now consider a beam with \( k = 2\pi/(632 \text{ nm}) \), and where the width of the lowest order coherent mode \( \sigma \) is approximately 5 mm to ensure paraxiality. The parameter \( \alpha \) characterizing the width of the exponential functions in equation (A.2) is given by \( \alpha = 1/2\sigma^2 = k/\nu \), so we set \( \nu = 497 \). Figure 2 shows the optimal coherence function \( A(x_1, x_2) \) (left column) and the optimal intensity distribution \( \mathcal{A}(x, x) \) (right column) at the \( z = 0 \) plane for the same values of \( \mu_2 \). We see that the larger \( \mu_2 \) is, the narrower the intensity distribution becomes. The shape of the coherence function changes from highly elongated to circular with increasing \( \mu_2 \), indicating that the width of \( \mathcal{A}(x_1, x_2) \) increases roughly as \( 1/\mu_2 \). Notice that, while always real, these correlations can take on negative values within small
regions off the diagonal. Surprisingly, these negative regions are only present for partially coherent fields, and they disappear in the coherent limit.

The results described above can be readily generalized to two transverse dimensions. In this case, for each representation, the total width is just the square root of the sum of the squares of the widths in each direction. If we assume that the profiles of the optimal solutions have rotational symmetry, then (analogously with the 1D calculation) it is easy to find that

$$\mu_2 = \sqrt{\frac{6 - 2(4 - g)gN + 6g^2N^2}{(N + 1)(N + 2)(3 - 4gN)^2}},$$

(33)

which for large $N$ gives $\mu_2 \approx 3/N$. The uncertainty relation for fields with specified $\mu_2$ is given by

$$\Delta_a\Delta_b \geq \frac{\eta^{(2D)}}{k},$$

(34)

where

$$\eta^{(2D)} = \frac{3 + (2 - 5g)N - 3gN^2}{3 - 4gN}.$$  

(35)

From these relations it readily follows that $\sqrt{3}/2 \leq \mu_2\eta^{(2D)} \leq 1$, so that the right-hand of equation (34) can be replaced with $\sqrt{3}/2k\mu_2$, i.e., it is inversely proportional to $\mu_2$ and not to $\mu_2^2$ as in the 1D case.

A case not considered explicitly here but worked out by Bastiaans [42] is that where the overall degree of coherence with $\bar{f}_c(B)$ of relation (26), based on Shannon’s entropy, is used. This leads to an infinite sum of modes that can be evaluated analytically, so that the optimal beam is of the Gaussian Schell-model form, in which $\mathcal{A}(x_1, x_2)$ is the product of a Gaussian in $x_1 + x_2$ and a Gaussian in $x_1 - x_2$.

Notice that results similar to those obtained above for 1D or 2D paraxial beams are found also in other situations which are physically different but in which the variables constitute a 1D or 2D Fourier-transform pair. In particular, for temporal pulses in linear dispersive media the
Figure 2. Illustration of the optimal coherence function (left column) and the optimal intensity distribution (right column) of a paraxial one-dimensional beam for various values of the overall degree of coherence: top row $\mu_2 = 0.2$, middle row $\mu_2 = 0.6$ and bottom row $\mu_2 = 1.0$. 
amplitude $A(p)$ is the pulse envelope $U(z = 0, \tau)$ with $p = \tau$ being the retarded time. The corresponding operators are $\hat{\mathcal{O}}_a = \tau^2$ and $\hat{\mathcal{O}}_b = -\beta_2^2 \frac{d^2}{d\tau^2}$ which lead to equation (29) with $k$ replaced by $1/\beta_2$. Another example is given by a 1D quantum particle for which $A(p)$ is the wave function and $p$ is the position coordinate $x$. In this case the operators are the squares of the position and momentum operators $\hat{\mathcal{O}}_a = x^2$ and $\hat{\mathcal{O}}_b = -\hbar^2 \frac{d^2}{dx^2}$, respectively, which result in the Schrödinger equation of the form of equation (29) where $1/\hbar$ substitutes $k$.

5.2. Pulse propagation under pure cubic dispersion

The second example is that of pulse propagation in an optical fiber with no linear dispersion. Such a situation is encountered, e.g., in fused silica at $\omega_0 = 1.483 \times 10^{15}$ Hz (known as the zero-dispersion frequency) corresponding to the wavelength $\lambda_0 = 1.27$ μm. We aim at finding the optimal spectral coherence function $\mathcal{A}(\omega_1, \omega_2)$ related to the initial field $A(\omega)$ at $z = 0$ for which the spectral width and its rate of increase under propagation are jointly minimized with $\mu_2$ fixed. The coherent situation was recently considered in [32].

The dispersion relation around the zero-dispersion frequency can be written as

$$k(\omega) = \beta_0 + \beta_1(\omega - \omega_0) + \frac{\beta_2(\omega - \omega_0)^3}{6},$$

where the dispersion coefficients are $\beta_0 = 7.16 \times 10^6$ m$^{-1}$, $\beta_1 = 4.88 \times 10^{-9}$ s m$^{-1}$, and $\beta_2 = 7.34 \times 10^{-41}$ s$^3$ m$^{-1}$ [32, 44]. The truncation at the cubic term was found to be sufficiently accurate within the interval $[0.4\omega_0, 1.6\omega_0]$ where the relative error in $k(\omega)$ is less than 0.5%. For the present case the operators are $\hat{\mathcal{O}}_a = -\frac{d^2}{d\omega^2}$ and $\hat{\mathcal{O}}_b = [k'(\omega) - k'(\omega_0)]^2$, resulting in the equation which is of the form of (29), but $V(\omega) = \nu^2 [k'(\omega) - \beta_1]^2$ and $\xi = \lambda$. Equation (29) therefore corresponds to a Shrödinger equation with quartic potential centered at $\omega_0$. Let us choose $\nu = 10^{-4}$, which as we will see yields a pulse with a reasonable spectral width, and as in [32] we set $\alpha = 70(\nu\beta_3)^{2/3}$. Following the previous example, we use the overall degree of coherence corresponding to $f_2(\hat{B})$ of equation (23). The procedure described in appendix A works well for a quartic potential since it is not too far from the quadratic one for which the Hermite–Gaussian functions are the eigenfunctions.

Figure 3 illustrates the $\tilde{b}_n$ coefficients in the cases of $\mu_2 = 0.15$, $\mu_2 = 0.6$, and $\mu_2 = 1.0$. Again, as expected, the more coherent the pulse is, the smaller is the number of modes that contribute; in the fully coherent case only the lowest order mode contributes. Figure 4 shows the spectral coherence function $\mathcal{A}(\omega_1, \omega_2)$ (left column) and the spectrum $\mathcal{A}(\omega, \omega)$ (right column) for the optimal pulse with the same three values of the overall spectral degree of coherence $\mu_2$ as in figure 3. We observe that when $\mu_2$ increases the spectrum gets narrower and its form approaches a shape which is known to be close to Gaussian [32]. In addition, the figure indicates that, as could be expected, the correlations in the spectral domain extend over larger frequency ranges when $\mu_2$ increases. Again, these functions are real, but they present negative regions, particularly for intermediate values of the overall level of coherence.

5.3. Fourier series relations and quantum phase

The third and last example deals with periodic functions and their conjugate Fourier series representations. There are several physical situations corresponding to this type of relation, such as trains of identical pulses generated, e.g., by a mode-locked laser (where the periodic variable
is time), or the diffraction of a monochromatic plane wave by a periodic grating. A third situation of this type is that of quantum phase, which is complementary to photon number. In this case the function $A(p)$ corresponds to the phase wavefunction $\psi(\theta)$ whose squared magnitude describes the probability density that a certain phase $\theta$ occurs, while the square magnitude of the Fourier coefficients of $\psi(\theta)$ are related to the (discrete) photon number probability distribution.

It must be mentioned that in the contexts of laser pulse trains and quantum phase, the Fourier coefficients corresponding to negative values of the index must vanish. In the case of pulse trains, this restriction is due to the use of the analytic signal representation of the field, while in the case of quantum phase it results from the fact that there cannot be negative photon numbers. For simplicity we disregard this restriction in what follows. Note, though, that doing so does not affect the validity of the lower bounds we find, since these are valid for any periodic function, whether or not its Fourier coefficients vanish for negative values of the index. That is, the resulting inequalities must be satisfied by physical solutions as well as by unphysical ones. Further, notice that we can factorize a linear phase factor by defining $A(\theta) = \psi(\theta) \exp(-i\bar{n}\theta)$, where $\bar{n}$ is the mean number of photons in the case of quantum phase. This way, the Fourier coefficients for $A$ can be nearly symmetric around $n = 0$ while those of $\psi$ vanish for $n < 0$, since the sequence is shifted by $\bar{n}$. The states that will be closest to the lower bounds obtained below will naturally be those for which $\bar{n}$ is much larger than $\Delta_n$ (the width of the distribution in $n$).

The task is to find the correlation function $A(\theta_1, \theta_2) = \langle A^*(\theta_1)A(\theta_2) \rangle$ such that the widths in $\theta$ and $n$ are jointly minimized when the overall degree of correlation $\mu_2$ is fixed. Given that we factored out a carrier phase proportional to $\bar{n}$, we use the centered operator $\hat{O}_n = -d^2/d\theta^2$ to estimate the width in $n$ of $A$. The operator used here to find the width in $\theta$ is the one proposed by Bandilla and Paul [11] within the context of quantum phase and used by many others [12–17] more generally for physical problems involving Fourier series or discrete Fourier transforms. Assuming that the distribution of the minimum uncertainty state is centered at $\theta = 0$, we choose the operator $\hat{O}_\theta = -\cos \theta$. Inserting the operators into equation (5) leads to the eigenvalue equation of (29) with $V(\theta) = -\nu^2 \cos \theta$ and $\xi = \lambda$. The eigenvectors are obtained in terms of
Figure 4. Illustration of the optimal spectral coherence function (left column) and the optimal spectrum (right column) of pulse propagation under pure cubic dispersion for several values of the overall degree of coherence: top row $\mu_2 = 0.15$, middle row $\mu_2 = 0.6$, and bottom row $\mu_2 = 1.0$. 
the even and odd Mathieu functions and the eigenvalues with the related characteristic values \[17\]. Let us again consider the case where \(\mu_2\) is used to characterize the purity of the field. The uncertainty relations corresponding to several values of this measure are represented graphically in figure 5(a): for a specified value of \(\mu_2\), the spreads \(\Delta_\theta\) and \(\Delta_n\) (mapped to \(\arctan(2\Delta_n)\) in the figure to make it compact) can never correspond to a point below the corresponding curve. Notice that these relations cannot be written in a simple analytical form as in the Fourier transform case in equation (31); instead they are found numerically and plotted parametrically for varying \(\nu\) through the procedure outlined earlier. The symbols along three of these curves \((\mu_2 = 0.2, \mu_2 = 0.6, \text{ and } \mu_2 = 1.0)\) correspond to \(\eta = 1\) for the specified values of \(\mu_2\). (b) Normalized coefficients \(\tilde{b}_n\) corresponding to these three optimal states.

6. Conclusions

As mentioned earlier, the analysis provided here applies not only to physical quantities that depend only on one continuous degree of freedom, but also to multidimensional (discrete and/or continuous) problems. Such cases would include, for example, nonparaxial electromagnetic fields, for which several measures of spatial and directional spread have been proposed \[23–29\], and where the effects of orbital and spin angular momentum have been considered \[30\].

![Figure 5](image-url)
Figure 6. Illustration of the optimal correlation function (left column) and the optimal normalized wavefunction (right column) for several values of the overall degree of phase correlation: top row $\mu_2 = 0.2$, middle row $\mu_2 = 0.6$, and bottom row $\mu_2 = 1.0$. 
Similarly, while in the last example we used the measure for the periodic variable employed in [11–17], one could also use alternative measures, such as those in [7, 8, 18].

As also discussed earlier, for the case of quantum phase versus photon number considered in the third example, the Fourier series is strictly constrained to have zero coefficients for \( n < 0 \). In the calculations presented in that example, we disregarded this restriction since our goal was to obtain lower bounds that are independent of \( \bar{n} \). However, the general method presented in this article can also be applied to find rigorous solutions in which the desired mean photon number is also specified, although such procedure would be more complicated. The key is to include in (3) a new term proportional to the functional derivative of \( \bar{n} \) (times another Lagrange multiplier) and then to write the resulting Sturm–Liouville equation in the \( n \) representation (the space of Fourier coefficients), so that it can be written in matrix form. This matrix is semi-infinite, reflecting the semi-infinite nature of the photon-number representation. The eigensystem can then be solved approximately by truncating the matrix at a sufficiently large \( n \) and using only the first few eigenvectors (which approximate the eigenvectors of the semi-infinite matrix). The new Lagrange multiplier must be varied such that resulting states have the prescribed \( \bar{n} \), which means that an iterative process might be required.

**Acknowledgments**

This work was partly funded by the National Science Foundation (PHY-1068325) and the Academy of Finland (projects 268705 and 268480). MAA thanks Aalto University in Espoo, Finland for the hospitality during his academic visit. TS and ATF also acknowledge Dean’s special support for coherence research at the University of Eastern Finland, Joensuu (project 930350).

**Appendix A. Numerical solution of equation (5)**

In the situations considered in this work one ends up with the eigenvalue equation

\[
-A''(p) + V(p - p_0)A(p) = \lambda A(p),
\]

whose eigenvectors and eigenvalues are denoted as \( \tilde{A}_n(p) \) and \( \tilde{\lambda}_n \), respectively. The expansion of \( A(p) \) in terms of Hermite–Gaussian functions is

\[
A(p) = \sum_{n=0}^{\infty} \frac{c_n}{\sqrt{\pi/\alpha}^{1/2}2^n n!} H_n\left( \sqrt{\alpha} (p - p_0) \right) \exp \left[ -\frac{\alpha(p - p_0)^2}{2} \right],
\]

where \( c_n \) and \( \alpha > 0 \) are constants, and \( H_n(x) \) is the Hermite polynomial of order \( n \). The functions \( H_n(x) \) satisfy the integral relations [45]

\[
\int_{-\infty}^{\infty} H_n(\sqrt{\alpha} x) H_m(\sqrt{\alpha} x) \exp \left( -\alpha x^2 \right) dx = \frac{\pi}{\sqrt{\alpha}} 2^n n! \delta_{n,m},
\]
\[
\int_{-\infty}^{\infty} \alpha x^2 H_n(\sqrt{\alpha} x) H_m(\sqrt{\alpha} x) \exp \left(-\alpha x^2\right) \, dx
= \frac{\sqrt{\pi}}{\alpha} \left[ 2^{n-1} (2n+1)n! \delta_{n,m} + 2^n (n+2)! \delta_{n+2,m} + 2^{n-2} n! \delta_{n-2,m} \right]
\] (A.4)

and the recurrence relations
\[ H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x), \] (A.5)
\[ H'_n(x) = 2n H_{n-1}(x). \] (A.6)

Substitution of expression (A.2) into (A.1) leads to the eigenvalue equation
\[
\sum_{n=0}^{\infty} H_{m,n} c_n = \lambda c_m,
\] (A.7)

for the coefficients \( c_m \), where
\[
H_{m,n} = \frac{\alpha}{2} (2n+1) \delta_{n,m} - \frac{\alpha}{2} \sqrt{(n+1)(n+2)} \delta_{n+2,m}
- \frac{\alpha}{2} \sqrt{n(n-1)} \delta_{n-2,m} + V_{n,m},
\] (A.8)

with
\[
V_{n,m} = \frac{1}{\sqrt{2(2n+1)!\pi n!}} \int_{-\infty}^{\infty} V(p) \frac{d^2}{dp^2} \left( H_n(p) H_m(p) \exp \left(-p^2\right) dp. \right)
\] (A.9)

References

[1] Cohen L 1995 *Time–Frequency Analysis* (New York: Prentice-Hall) chapter 3
[2] Folland G and Sitaram A 1997 The uncertainty principle: a mathematical survey *J. Fourier Anal. Appl.* 3 207–38
[3] de Bruijn N G 1967 Uncertainty principles in Fourier analysis *Inequalities* ed O Shisha (San Diego, CA: Academic) pp 57–71
[4] Heisenberg W 1927 Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik *Z. Phys.* 43 172–98
[5] Siegman A E 1990 New developments in laser resonators *Proc. SPIE* 1224
[6] Jackiw R 1968 Minimum uncertainty product, number-phase uncertainty product, and coherent states *J. Math. Phys.* 9 339–46
[7] Pegg D T and Barnett S M 1989 Phase properties of the quantized single-mode electromagnetic field *Phys. Rev. A* 39 1665–75
[8] Pegg D T and Barnett S M 1997 Tutorial review: quantum optical phase *J. Mod. Opt.* 44 225–64
[9] Lynch R 1995 The quantum phase problem: a critical review *Phys. Rep.* 256 367–436
[10] Bialynicki-Birula I, Freyberger M and Schleich W 1993 Various measures of quantum phase uncertainty: a comparative study *Phys. Scr.* T48 113–8
[11] Bandilla A and Paul H 1969 Laser amplifier and phase uncertainty *Ann. Phys., Lpz.* 23 323–33
[12] Lukš A and Peřinová V 1991 Extended number state basis and number-phase intelligent states of light fields: I. Mapping and operator ordering approach to quantum phase problems *Czech. J. Phys.* 41 1205–30
[13] Opatrný T 1994 Mean value and uncertainty in optical phase—a simple mechanical analogy J. Phys. A: Math. Gen. 27 7201–8
[14] Opatrný T 1995 Number-phase uncertainty relations J. Phys. A: Math. Gen. 28 6961–75
[15] Forbes G W and Alonso M A 2001 Measures and spread for periodic distributions and the associated uncertainty relations Am. J. Phys. 69 340–7
[16] Forbes G W and Alonso M A 2001 Consistent analogs of the Fourier uncertainty relation Am. J. Phys. 69 1091–5
[17] Forbes G W, Alonso M A and Siegman A E 2003 Uncertainty relations and minimum-uncertainty states for the discrete Fourier transform and the Fourier series J. Phys. A: Math. Gen. 36 7027–47
[18] Alonso M A and Bastiaans M J 2004 Mapping-based width measures and uncertainty relations for periodic functions Signal Process. 84 2425–35
[19] Franke-Arnold S, Barnett S M, Yao E, Leach J, Courtial J and Padgett M 2004 Uncertainty principle for angular position and angular momentum New J. Phys. 6 103–10
[20] Pegg D T, Barnett S M, Zambrini R, Franke-Arnold S, Yao E and Padgett M 2005 Minimum uncertainty states of angular momentum and angular position New J. Phys. 7 62–81
[21] Yao E, Franke-Arnold S, Courtial J, Barnett S M and Padgett M 2006 Fourier relation between angular position and optical orbital angular momentum Opt. Express 14 9071–6
[22] Hradil Z, Řeháček J, Bouchal Z, Čelechovský R and Sanchez-Soto L 2006 Minimum uncertainty measurements of angle and angular momentum Phys. Rev. Lett. 97 243601
[23] Porras M A 1994 The best quality optical beam beyond the paraxial approximation Opt. Commun. 111 338–49
[24] Alonso M A and Forbes G W 2000 Uncertainty products for nonparaxial wave fields J. Opt. Soc. Am. A 17 2391–402
[25] Sheppard C J R 2001 High-aperture beams J. Opt. Soc. Am. A 18 1579–87
[26] Alonso M A, Borghi R and Santarsiero M 2006 Nonparaxial fields with maximum joint spatial-directional localization: I. Scalar case J. Opt. Soc. Am. A 23 691–700
[27] Alonso M A, Borghi R and Santarsiero M 2006 Nonparaxial fields with maximum joint spatial-directional localization: II. Vectorial case J. Opt. Soc. Am. A 23 701–12
[28] Sheppard C J R, Alonso M A and Moore N J 2008 Localization measures for high-aperture wavefields based on pupil moments J. Opt. A: Pure Appl. Opt. 10 033001
[29] Bialynicki-Birula I and Bialynicka-Birula Z 2013 Uncertainty relation for focal spots in light beams Phys. Rev. A 88 052103
[30] Alonso M A 2011 The effect of orbital angular momentum and helicity in the uncertainty-type relations between focal spot size and angular spread J. Opt. 13 064016
[31] Ponomarengo S A and Agrawal G P 2008 Phase-space quality factor for ultrashort pulsed beams Opt. Lett. 3 767–9
[32] Alonso M A, Setälä T and Friberg A T 2011 Optimal pulses for arbitrary dispersive media J. Eur. Opt. Soc. Rapid Publ. 6 11035
[33] Sakurai J J and Napolitano J 2011 Modern Quantum Mechanics 2nd edn (Boston: Pearson) pp 181–3
[34] Born M and Wolf E 2002 Principles of Optics 7th edn (Cambridge: Cambridge University) ch 10
[35] Mandel L and Wolf E 1995 Optical Coherence and Quantum Optics (Cambridge: Cambridge University)
[36] Wolf E 2007 Introduction to the Theory of Coherence and Polarization of Light (Cambridge: Cambridge University)
[37] Schmidt E 1907 Zur Theorie der linearen und nicht-linearen Integralgleichungen Math. Ann. 63 433–76
[38] Bastiaans M J 1982 Uncertainty principle for spatially quasi-stationary, partially coherent light J. Opt. Soc. Am. 72 1441–3
[39] Bastiaans M J 1983 Uncertainty principle for partially coherent light J. Opt. Soc. Am. 73 251–5
[40] Bastiaans M J 1983 Lower bound in the uncertainty principle for partially coherent light J. Opt. Soc. Am. A 1 1320–4
[41] Bastiaans M J 1984 New class of uncertainty relations for partially coherent light J. Opt. Soc. Am. 72 711–5
[42] Bastiaans M J 1986 Uncertainty principle and informational entropy for partially coherent light J. Opt. Soc. Am. A 3 1243–6
[43] Blomstedt K, Setälä T and Friberg A T 2007 Effective degree of coherence: general theory and applications to electromagnetic fields J. Opt. A: Pure Appl. Opt. 9 907–19
[44] Agrawal G P 1995 Nonlinear Fiber Optics 2nd edn (San Diego, CA: Academic)
[45] Arfken G B and Weber H J 2005 Mathematical Methods for Physicists 6th edn (Amsterdam: Elsevier)