Modal Analysis of Partially-Coherent Submillimetre-Wave Quasioptical Systems

S. Withington ¹ and J.A. Murphy²,
¹Cavendish Laboratory, University of Cambridge, UK
²St. Patrick's College, Maynooth, Ireland

March 19, 1997

Abstract: We consider the modal analysis of partially-coherent submillimetre-wave quasioptical systems. According to our scheme the cross-spectral density is expanded as a sum of partially-coherent propagating free-space modes. The coherence matrix, the elements of which are determined by evaluating bimodal overlap integrals, completely describes the state of the field at a plane and can be traced through the optical system to another plane by means of a scattering matrix. Whereas diagonalising the scattering matrix gives the natural modes of the optical system, diagonalising the coherence matrix gives the natural modes of the field. As a special case, we consider the case where the field at the source plane is completely incoherent. After developing a number of analytical tools, we demonstrate the technique by analysing the behaviour of a Gaussian-beam telescope. Throughout the paper we emphasise the physical significance of the equations derived.

1 Introduction

We consider the Gaussian-mode analysis of partially-coherent submillimetre-wave quasioptical systems. Our primary aim is to show that the modal techniques that have been developed for the analysis of coherent optics [1, 2, 3] can be extended to cover the case when the field propagating through the system is partially coherent. Partial coherence arises in some form in all problems of practical importance. For example, consider the case where a submillimetre-wave telescope is used to observe an extended radio astronomical source. Usually, for heterodyne receivers, one would calculate the coupling between the detector and the source by propagating the fully-coherent field of the horn 'backwards' through the optical system onto the sky [4, 5]. There is no reason why, however, the analysis should not proceed in the opposite direction; that is to say it should be possible, at least in principle, to propagate the fully-incoherent field of the source 'forward' through the optical system onto the focal plane. In the first case, the field passing through the system is single moded even though the optical system itself is multimoded; whereas, in the second case, the field is multimoded. Clearly, one has to distinguish between the modal properties of the field and the modal properties of the optical system through which the field is passing.

In this paper, we describe a procedure that allows the second-order statistical properties of a field to be traced through a complex system of long-wavelength optical components. The partially-coherent field is constructed from coherent, diffracting free-space modes. A key feature of the scheme is that the description of the field is complete—in addition to the description of the optical system being complete—and therefore problems of almost any complexity can be solved. For example, we are currently investigating the behaviour multi-mode bolometer imaging arrays, where the state of the overall field is described by a single coherence matrix. Even in the case of a single-
mode detector, the unused, orthogonal modes will be excited by noise, and this noise appears in the analysis in a natural and elegant manner. Not only does the technique allow the throughput of complex systems to be determined, but it also provides a considerable amount of physical insight into the way partially-coherent submillimetre-wave optical systems behave. Looking to the future, it should be possible to study the injection of noise by lossy scattering components and the effects of a fluctuating medium such as the atmosphere. At a deeper level, modal analysis provides an understanding of the thermodynamic entropy of a beam in a such a way that maximum entropy techniques could be used to examine, in some detail, the state of a collimated field from intensity measurements alone.

In the first part of the paper, we review the integral-equation description of the propagation of a partially-coherent field. This form provides the starting point for an analysis in terms of free-space modes. We then describe the way in which the second-order statistical properties of a free-space beam can be decomposed into a sum of partially-coherent modes which propagate easily and which can be scattered at optical components. The form of the partially-coherent field at a plane is fully characterised by a coherence matrix. We show that the coherence matrix can be diagonalised to render the natural modes of the field. These modes are individually fully coherent but have no well-defined phase relationship between each other. In this sense they propagate in an independent manner.

We explain how the elements of the coherence matrix can be determined through bimodal overlap integrals, and we show how the mode set should be chosen so that the field can be represented with near-optimum numerical efficiency. We describe how the coherence matrix can be propagated in both the forward and backward directions. We then consider the special case when the illuminating field is fully incoherent. This situation occurs, for example, when the beam from a radio telescope and all of its sidelobes are coupled into an extended isothermal source. In this case coherence builds up as the field propagates and the natural modes of the field at the output plane are the same as the natural modes of the optical system.

To illustrate the technique we consider in some detail the one-dimensional Gaussian beam telescope. This particular example allows us to demonstrate clearly the basic physical concepts. We assume scalar fields, but this is not an intrinsic limitation of the scheme. The Gaussian beam telescope is analysed in terms of Gaussian-Hermite modes and recursion relationships are used to calculate the scattering matrices of the individual apertures. The behaviour of the whole system is then characterised by a single low-order scattering matrix which is simply the product of the scattering matrices of the individual components. When considering a more complicated system all that needs to be done is to replace the scattering matrix of the Gaussian-beam telescope with that of the actual system under consideration. To complete the paper, we present various useful analytical tools, and demonstrate the overall method by analysing the behaviour of a Gaussian-beam telescope when a partially-coherent field is applied.

2 Classical analysis

We assume that the submillimetre-wave system to be analysed comprises a sequence of components that interact with and scatter a propagating free-space beam [6, 7]. From a classical point of view, assuming for a moment full coherence, the optical system maps the field at the input plane onto the field at the output plane in a linear manner: $E_1(r_1) \mapsto E_2(r_2)$. Here, $r_1$ and $r_2$ represent position vectors on the input and output surfaces respectively, and a subscript on a quantity denotes the plane over which the quantity is being considered. Because the mapping is linear, we can express
the output field in the form
\[ E_2(r_2) = \int_{S_1} E_1(r_1) K(r_2|r_1) \, dS_1, \tag{1} \]
where the integral is evaluated over the input surface \( S_1 \). Although, it is tempting to assume that the input surface is flat and perpendicular to the beam, this does not have to be the case, and any surface which contains the beam can be used if the appropriate kernel \( K(r_2|r_1) \) is known. It is important to appreciate that the kernel is a function of both the input and output coordinates, and therefore the system does not have to be isoplanatic. As a consequence the formalism is applicable in cases where significant aberrations are present. This feature contrasts with Fourier Optics where the kernel has to be space invariant [8].

When cast in its modal form, the above equation can be used to analyse in detail the behaviour of fully-coherent submillimetre-wave optical systems [2]. In many cases, however, the source is not coherent and another level of sophistication is required. If we define the cross-spectral density of an ensemble of random fields to be
\[ W_2(r_2', r_2) = \langle E_2(r_2) E_2^*(r_2') \rangle, \tag{2} \]
where \( \langle \cdot \rangle \) denotes the ensemble average, and quasimonochromatic fields are assumed, then through (1) it is trivial to show that the cross-spectral density propagates according to
\[ W_2(r_2', r_2) = \int_{S_1} W_1(r_1', r_1) K(r_2| r_1) K^*(r_2'| r_1') \, dS_1 dS_1', \tag{3} \]
where \( r_1 \) and \( r_1' \) represent two different points on the input surface. Hence, we can calculate the cross-spectral density on the output surface if we know the cross-spectral density on the input surface. Again it is emphasised that the optical system does not have to be isoplanatic.

A special case arises when the field on the input surface is everywhere fully incoherent. This situation will only occur at the position of a source as any propagation or scattering will induce some degree of coherence, as will be shown later. In the case when the input is fully incoherent we can write
\[ W_1(r_1', r_1) = I(r_1) \delta(r_1 - r_1'), \tag{4} \]
where \( I(r_1) \) is the intensity of the source. Substituting in the expression for the propagation of the cross-spectral density, (3), we get
\[ W_2(r_2', r_2) = \int_{S_1} I_1(r_1') K(r_2| r_1') K^*(r_2'| r_1') \, dS_1'. \tag{5} \]
In general, an incoherent source will lead to coherence in the output plane.

Finally, if we are only interested in the intensity of the field at the output plane,
\[ I(r_2) = W_2(r_2, r_2) = \int_{S_1} I_1(r_1') |K(r_2| r_1')|^2 \, dS_1'. \tag{6} \]
We can see that for an incoherent source, and for the case where one is only interested in intensity, the output is linear in intensity and the kernel is the square modulus of the kernel associated with the fully-coherent case [9]. This equation describes, of course, the way in which a telescope images the brightness distribution of an astronomical source onto its focal plane.

The imaging properties of optical systems are usually understood in terms of Fourier expansions of the above equations. Unfortunately, Fourier Optics can only be used in the case of ideal, isoplanatic systems, and moreover the equations derived only relate fields at conjugate Fourier planes. What we would like is a simple scheme based on multimode Gaussian optics that allows the statistical properties of a field to be calculated at any intermediate plane.
3 Gaussian-mode expansion of the cross-spectral density

For partially-coherent systems, the quantity we wish to decompose into modes is the cross-spectral density: where, formally, we understand the cross-spectral density to be the time Fourier transform of the mutual coherence function. We can achieve this end by assuming that the optical system under consideration is one member of an ensemble. If the bandwidth is sufficiently narrow so that the coherence length is very much greater than the physical size of the system, the phase at one point in one member of the ensemble is well defined with respect to the phase at another point in the same member of the ensemble. As in the coherent case, the field can then be written as a modal sum [10]:

$$E_i(r, \omega) = \sum_m A_{m}^{i}(\omega) \psi_m(r, \omega),$$  \hspace{1cm} (7)

where we tacitly understand that if the system is two dimensional, the index labelling the mode represents two indices. In this equation, \(i\) denotes a particular member of the ensemble, and the frequency dependence of the mode coefficients has been indicated explicitly by \(\omega\). Obviously, we can now represent the cross-spectral density at some plane in terms of this expansion:

$$W_i(r_1, r) = \langle E_i(r_1) E_i^{*}(r) \rangle = \sum_m \sum_{m'} C_{m,m'} \psi_{m'}^{*}(r_1) \psi_m(r),$$  \hspace{1cm} (8)

where

$$C_{m,m'} = \langle A_{m}^{*} A_{m'}^{i} \rangle.$$  \hspace{1cm} (9)

The above equation is simply the bimodal expansion of the cross-spectral density. From a physical point of view, the non-negative-definite Hermitian form ensures that a positive intensity is formed when two parts of the field are combined in an interferometer. In the above equation we have dropped the explicit reference to frequency, but it must be remembered that to get the full behaviour over a range of frequencies, the appropriate integration must be carried out.

For convenience we can write the coefficients of the bimodal expansion in matrix form:

$$C = \langle A^i A^{i*} \rangle$$  \hspace{1cm} (10)

where \(A^i\) is the column vector of mode coefficients corresponding to the \(i\)'th member of the ensemble, and \(A^{i*}\) denotes the conjugate transpose. Clearly, once we know the expansion coefficients \(C\) we have characterised the form of the partially-coherent beam at a plane. The propagation of the coherence matrix in the partially-coherent case is equivalent to the propagation of the mode coefficients in the fully-coherent case.

Although physically appealing the above argument is not mathematically rigorous. The problem lies in the fact that, in order to generate the modal expansions, we tacitly assumed that we could Fourier transform the time-dependent field of each member of the ensemble, and yet it is well known that it is not possible, because of lack of convergence, to Fourier transform the members of a stationary random process. Wolf has considered the modal expansion of three-dimensional stationary random fields in some detail [11, 12]. He showed that in order to avoid the introduction of generalised Fourier transforms, it is possible to set up an ensemble which generates the cross-spectral density as an ensemble average of mode coefficients of ordinary functions. The arguments are somewhat involved, but the outcome is that modal expansions of the above kind are rigorously correct, as physical intuition would suggest. The essential point to bear in mind is that one is propagating a statistical property of the field rather than the field itself, and this quantity, the cross-spectral density, propagates according to Helmholtz equations.
We now need to know how to calculate the mode coefficients when the functional form of the cross-spectral density is known. By a simple extension of the usual overlap integral we find that the elements of the coherence matrix are given by

$$ C_{m,m'} = \int_{S_1} W(r_1', r_1) \psi_m^*(r_1') \psi_{m'}(r_1) dS_1 dS_{1'} . $$  \hspace{1cm} (11)

In general, we do not know the cross-spectral density at every plane, but there is usually some plane over which the cross-spectral density is known.

4 Natural modes of a partially-coherent field

At this stage we have expanded the cross-spectral density in terms of some basis mode set. We can of course transform to some other basis set and describe the field equally well. It is a basic feature of the Hermitian form given above, that if the cross-spectral density is to remain unchanged under this transformation, the transformation must be unitary. That is to say there is some transformation $U$, for which $UU^T = I$ and $UCU^T$ is diagonal. Of particular interest is the transformation that diagonalises the coherence matrix. In a sense, the modes found in this way are the natural modes of the field, because then the partially-coherent field is represented by a sum of modes which are individually fully spatially coherent but completely incoherent with respect to each other.

The problem of how to expand the cross-spectral density of a three-dimensional random field in terms of its natural modes has been considered by Wolf [11, 12], who approached the problem not by thinking about the system in which the field is contained but by thinking about the intrinsic properties of the field itself. The solution is to remember Mercer’s theorem, which essentially states that if the kernel of a homogeneous Fredholm equation of the second kind is Hermitian and nonnegative definite then the eigenvalue spectrum is discrete, real and the eigenvectors form a complete orthonormal set in terms of which the kernel can be expanded. When looking for a bimodal expansion, it is therefore natural to set up an integral equation of the form

$$ \lambda_i \phi_i (r_1) = \int W(r_1', r_1) \phi_i (r_1') dr_1' $$  \hspace{1cm} (12)

so that the kernel can be expressed as a weighted sum of eigenfunctions:

$$ W(r_1', r_1) = \sum_i \lambda_i \phi_i^*(r_1') \phi_i (r_1) . $$  \hspace{1cm} (13)

These eigenfunctions, $\phi_i$, are the natural modes of the field, and $\lambda_i$ are the associated eigenvalues.

We can now ask what is the relationship between Wolf’s natural mode set and the modes that we have used. If we expand the natural modes as a sum of our—as yet undefined—propagating modes, we have

$$ \phi_i (r_1) = \sum_n A_i^* \psi_n (r_1) , $$  \hspace{1cm} (14)

and if we substitute this equation together with the bimodal expansion (8) into the eigenvalue expression (12), we find that

$$ [C - \lambda_i I] A_i^* = 0 , $$  \hspace{1cm} (15)

where now $A_i^*$ are the mode coefficients of the natural mode $i$ and $\lambda_i$ is the associated eigenvalue. Hence, if we know the cross-spectral density at some plane, we can construct the coherence matrix by evaluating the overlap integrals and then we can find the natural modes by finding the eigenvectors through (15).

To understand the physical meaning of these expansions, suppose that we have an ensemble of optical systems, where the field associated with each member of the ensemble is constructed from
spatially-coherent modes, the phases of which are fully incoherent and uniformly distributed with respect to each other. In this case we have

$$C_{m,m'} = \left( |A_{m'}^*| |A_m^*| \right) \left( \exp \left[ j \left( \theta_m - \theta_{m'} \right) \right] \right) = \lambda_m \delta_{m,m'} \, .$$  

The coherence matrix is diagonal as one would expect. If, however, we now transform to another arbitrary mode set, it is easy to show and physically reasonable, that the mode coefficients become partially coherent and the coherence matrix is no longer diagonal.

In summary, we can expand the cross-spectral density in terms of any convenient mode set by using the bimodal form of the overlap integral. Because the modes are not chosen in any particular way, correlations will exist between the mode coefficients and the coherence matrix will be full. Here we are expanding the cross-spectral density in terms of a set of fully spatially-coherent modes which are partially coherent with respect to each other. If we diagonalise the coherence matrix, we can express the cross-spectral density as a sum of modes which are completely uncorrelated with respect to each other, and therefore have no definite phase relationship between them. These are the natural modes of the optical field defined by Wolf.

Expanding a random process as a sum of orthogonal uncorrelated functions is known as a Karhunen-Loève expansion, and in general an expansion can be found even when the process is not stationary. For example, in adaptive optics, the randomly distorted phase in the aperture of a mirror is usually expanded in terms of Zernike polynomials. The most efficient functions to use, however, when compensating for phase errors is the Karhunen-Loève expansion [13]. There is a close relationship between the technique that we are promoting and the techniques of adaptive optics, but it must be appreciated the physical application is very different. In this paper, we are expanding a propagating field, whereas in adaptive optics it is the phase of the field at a plane that is being expanded. Nevertheless it would be particularly interesting to use partially-coherent Gaussian modes to model the behaviour of a submillimetre-wave telescope when a turbulent atmosphere is included.

5 Completely-incoherent sources

Of particular significance is the case where the field at the input plane is fully incoherent and has uniform intensity. This situation occurs for example when the beam of a submillimetre-wave telescope is coupled to a source of uniform brightness. For an incoherent field, we can write

$$W(r_1,r_2) = I_0 \delta(r_1 - r_2) \, .$$  

Substituting this form into the bimodal overlap integral we find

$$C_{m,m'} = \int_{S_1} I(r_1) \psi_{m'}^*(r_1) \psi_m(r_1) \, dS_1 \, .$$  

Now if the source has uniform brightness

$$C_{m,m'} = I_0 \delta_{m,m'} \, ;$$

that is to say all of the modes are excited equally and independently. It can be shown that because of completeness this statement must be true regardless of what mode set is used. Physically, this must be true because the resultant field cannot contain any spatial information. In matrix form, we have for a uniform incoherent source

$$C = I_0 I$$

where I is the identity matrix. Obviously, if the brightness over the plane of the source is not uniform, correlations must be induced between modes even though the source itself is incoherent. This behaviour is to be expected classically because the van Cittert-Zernike theorem tells us that correlations exist in the far field of a source of finite size even when the source itself is incoherent.
6 Propagating the correlation matrix

It is well known that a coherent field can be traced through a submillimetre-wave optical system by multiplying the mode coefficients of the incoming beam by a scattering matrix. Moreover, if the optical system comprises a number of optical components then the scattering matrix associated with the overall system is simply the product of the scattering matrices associated with the individual components [2]. This procedure is fundamentally based on the modal expansion of equation (1). That is to say the input and output fields, and the kernel, are expanded in terms of a convenient set of modes prior to evaluating the integral. In the case of a partially-coherent beam, we need to find the modal equivalent of equation (3); or in other words, we need to ask how the coherence matrix can be traced through an optical system once the scattering matrix is known.

We know that for each member of the ensemble we can propagate the field according to the usual Gaussian mode scattering matrix \( S \). Hence if the field at the input plane has mode coefficients \( A_i \) then the field at the output plane has mode coefficients \( B_i \) where

\[
B_i = S A_i .
\]

The coherence matrix at the output plane therefore becomes

\[
\langle B_i B_i^T \rangle = S \langle A_i A_i^T \rangle S^T ;
\]

or

\[
D = S C S^T ,
\]

where \( D \) is the coherence matrix at the output plane, \( S \) is the usual coherent-mode scattering matrix, and \( C \) is the coherence matrix at the input plane. Hence we can calculate the scattering matrix associated with the optical system in the usual way, and then calculate the coherence matrix at the output plane if we know the coherence matrix at the input plane. Numerically, the procedure is very straightforward, and the scattering matrix only has to be calculated once for a given optical system. The scattering matrix contains all of the information necessary to propagate a coherent or incoherent field. Notice also that if the optical system under consideration is varying with time, as would be the case for a turbulent atmosphere above a telescope, then the ensemble average should include the time-varying scattering matrix.

It is also interesting to ask how we calculate the cross-spectral density at the input plane if we know the cross-spectral density at the output plane. This is clearly the imaging process in the case where we have complete knowledge about the amplitude and phase of the cross-spectral density in the focal plane. It is straightforward to show that as long as \( S^{-1} \) exists, the cross-spectral density at the input plane is given by

\[
C = S^{-1} D (S^{-1})^T .
\]

In a later paper, we will show how this leads to a general method for reconstructing images even in the case when the imaging array is far from ideal.

7 Analysis of partially-coherent submillimetre-wave optical systems

Before demonstrating the above techniques through the analysis of a particular system, it is useful to review the overall procedure with an emphasis on physical interpretation.

We now know that the properties of a partially-coherent field can be fully characterised by means of a coherence matrix. The coherence matrix is simply a convenient way of organising the
coefficients of a bimodal expansion of the cross-spectral density. The modes used in the expansion can be any orthonormal set, but it is convenient to use a set of modes that propagate easily. Later we will show that Gaussian-Hermite or Gaussian-Laguerre modes are almost ideal. Once the coherence matrix is known, it can be traced forwards or backwards through the optical system by using the ordinary coherent-mode scattering matrix. Finally, once the scattered coherence matrix is known, the cross-spectral density can be reconstructed at the output plane.

It is also straightforward to show that if we combine two fields at a plane that are generated incoherently then the overall coherence matrix is the sum of the two individual coherence matrices regardless of the states of coherence of the individual fields. This particular theory has application for example when a submillimetre-wave beam is truncated by a lossy aperture which injects noise of its own. We will show in a later paper how to calculate the coherence matrix of a lossy, passive component simply from knowledge of its temperature and scattering matrix. In general, therefore, not only can we propagate a partially-coherent field but we can also add in noise generated by lossy components. If required these noise sources can be referenced to one end of the optical system giving a set of noise parameters which completely characterises the noise performance of the system in much the same way as the noise properties of a microwave transistor are characterised by a set of noise parameters which are referenced to the input. In the case of an optical system the noise parameters take the form of a matrix of complex temperatures.

It is vitally important to appreciate that the coherence matrix characterises the modal properties of the field at a plane whereas the scattering matrix completely characterises the modal properties of the optical system. The two are, of course, distinct. In general, because an arbitrary mode set has been chosen to expand the cross-spectral density, the coherence matrix will be full showing that correlations exist between the modes. We can, however, diagonalise the coherence matrix to give the mode coefficients of the natural modes of the field, that is to say the modes that propagate independently with full spatial coherence but no definite phase relationship between them. The eigenvalues give the amount of power in each mode. In general, even though we do not diagonalise the coherence matrix we should choose a convenient mode set that makes the coherence matrix as near diagonal as possible. In this way the modes chosen will be as close as possible to the true natural modes of the field.

In a previous paper [2], we discussed the diagonalisation of the scattering matrix. In this case the eigenvectors give the mode coefficients of field distributions which pass through the optical system unchanged, and the eigenvalues give the loss associated with the propagation of these fields. These modes are the normal modes of the optical system, and in general any incoming field distribution can be expanded in terms of these modes and propagated through the system simply by multiplying by the eigenvalues. For convenience we choose a mode set that propagates easily and yet which diagonalises as near as possible the scattering matrix. In general the natural modes of the optical system are not the same as the natural modes of the field, and we can ask whether it is more reasonable to choose a mode set that near diagonalises the coherence matrix or a mode set that near diagonalises the scattering matrix.

It is now particularly revealing to ask what happens in the case where an optical system described by a scattering matrix $S$ is illuminated by a uniform fully-incoherent source. We know that for any mode set the coherence matrix of a fully-incoherent source is diagonal, and therefore the field at the output plane of the optical system is described by

$$D = \mathbf{I}_s \mathbf{S} \mathbf{I}_s^{-T}.$$  \hspace{1cm} (25)

It is clear that although the field at the input plane is fully incoherent, the field at the output plane has coherence induced on it due to mode filtering. The induction of coherence is evidenced by the appearance of off-diagonal terms in the output coherence matrix. We know, however, that there is some mode set that diagonalises the scattering matrix—the natural modes of the optical
system—and clearly in this case the coherence matrix at the output plane must also be diagonal. Hence, if the optical system is illuminated by an incoherent field, the natural modes of the field at some plane in the optical system are the same as the natural modes of the optical system itself. This statement seems physically reasonable because all of the coherence induced in the originally incoherent field is due to the optical system.

It is also interesting to note at this stage that spatial coherence will generally build up as modes are filtered. We know that when a low-throughput optical system is illuminated by a coherent source, the output of the system will tend to the lowest-order eigenmode regardless of the precise nature of the incoming field, and this is why submillimetre-wave optical systems always, somewhat conveniently, tend to produce Gaussian beams. In the case of incoherent illumination the situation is somewhat similar, but now as more and more low-throughput components are added, the output field will tend to become more and more coherent. We will demonstrate this effect in the next section.

In summary, it is important to distinguish between the natural modes of the optical system and the natural modes of the field. The natural modes of the optical system are found by diagonalising the scattering matrix whereas as the natural modes of the field are found by diagonalising the coherence matrix. In the case where the incoming field is incoherent the two are, as shown above, identical. If the incoming field is partially coherent, however, the natural modes of the field after passing through the optical system will not be the same as those of the optical system, and the similarities will depend on the degree to which the optical system imposes coherence on the field.

A further important and useful consideration is that the number of significant non-zero eigenvalues found when diagonalising the scattering matrix gives the number of degrees of freedom of the optical system. The number of non-zero eigenvalues found when diagonalising the coherence matrix gives the number of degrees of freedom of the field. Clearly, the number of degrees of freedom of the field can only be as many as the optical system and the two will be the same when the incident field is fully incoherent. In the case where the incoming field is fully coherent, the coherence matrix at any plane will have only one non-zero eigenvalue, a feature which can be traced to the fact that the elements of the coherence matrix factorise. In general, this will mean that because the coherence matrix has \(N^2\) elements it will not be possible to diagonalise. The solution to this apparent paradox is that the modes of the optical system will all, at some level, be excited by noise. In fact it can easily be seen that we can add a noise coherence matrix to a fully-coherent coherence matrix we get a matrix that can be diagonalised, and if the signal to noise ratio is high, one eigenvalue will be much greater than all of the others. The ability to study the signal and noise properties of a field through diagonalisation is a particularly powerful technique. For example, the field 'produced' by a complete array of detectors \([14]\) can be described by a single coherence matrix, the eigenvalues of which give the relative sensitivities of the pixels. Another example is the thermal radiation emitted by an overmoded horn when a perfect absorber is placed in the overmoded waveguide. Moreover, the diagonalisation of the coherence matrix could be used to diagnose problems in experimentally-derived data. In particular it should be possible to extract the coherent field of interest from background noise. Before leaving the subject of modes, it is also worth pointing out that the coherence matrix just described is the optical analogue of the quantum mechanical density matrix, where a coherent field corresponds to a pure quantum state and a partially-coherent field corresponds to a mixed quantum state.

In addition to diagonalising the coherence matrix for the purpose of investigating the nature of a field, we can also use various other analytical tools. First it should be noticed that the total amount of power in the beam is given by the trace of the coherence matrix:

\[
P = \sum_i C_{i,i} = Tr(C).
\]

Now the trace of a matrix is invariant to unitary transformations such as diagonalisation, and
therefore, the total amount of power is also given by the sum of the eigenvalues (the eigenvalues are real because the coherence matrix is Hermitian). Physically, this is to be expected because the diagonalised coherence matrix is representing the cross-spectral density as a sum of completely incoherent modes. Moreover, although we will not discuss it in this paper it is likely that the eigenvalues can be used as a measure of the degree of disorder of the whole field, and therefore it should be possible to use maximum entropy techniques to recover mode coefficients from noisy partially-coherent experimental data [15, 16].

Once the coherence matrix is known, we can reconstruct the intensity and degree of coherence of the field. As stated earlier the cross-spectral density is given by

$$W(r', r) = \sum_m \sum_{m'} C_{m,m'} \psi_m^*(r_1) \psi_m(r_1), \tag{27}$$

and therefore the intensity and degree of coherence can be written

$$I(r) = W(r, r) \tag{28}$$

and

$$\Gamma(r', r) = \frac{W(r', r)}{\sqrt{I(r')I(r)}} \tag{29}$$

respectively. The one-dimensional forms of these expressions will be used in the next section.

8 Partial coherence and the Gaussian-beam telescope

In the preceding sections we outlined a method by which the behaviour of multimode partially-coherent submillimetre-wave quasioptical systems can be analysed. Although the technique can be used to analyse the behaviour of almost any system—through the use of the appropriate scattering matrix—in this section we shall apply the theory to the Gaussian-beam telescope [1]. Not only does the Gaussian-beam telescope exhibit features which are integral to all systems, but it also produces results which are easily interpreted in terms of classical analysis [17, 18]. In order not to cloud the central features of the model, we shall work in one dimension, but the extension to two dimensions is straightforward.

A diagram of a Gaussian-beam telescope is shown in Fig. 1. This arrangement is important because the field at the input plane is imaged onto the field at the output plane in a frequency-independent way. An important feature of the arrangement is that there are two apertures, one at the input plane and one at the conjugate Fourier plane. In the context of a telescope, one can be regarded as the aperture stop and one as the field stop. Although for convenience we have located
the apertures at particular planes, the technique can easily handle other arrangements. The reason for including apertures in the model is that they limit the throughput of this ideal imaging system as would be the case for any real system with finite-size components.

To begin, we must calculate the scattering matrix of the system. Let us for the moment assume that the fields in the regions between the components are described as sums of propagating Gaussian-Hermite modes. Each mode has the form

$$
\psi_m(z) = \left( \frac{\sqrt{2}}{w} \right)^{1/2} h_m \left( \frac{\sqrt{2} \pi}{w} \right) \exp \left[ \pm j \theta \right] \exp \left[ \mp j \frac{\pi z^2}{\lambda R} \right] \exp \left[ \mp j kz \right]
$$

(30)

where

$$
\psi_m(u) = \frac{H_m(u) \exp \left[ -\frac{u^2}{2} \right]}{\left( 2^m m! \right)^{1/2}},
$$

(31)

and

$$
\theta = (m + 1/2) \frac{z}{z_e},
$$

(32)

also $H_m(u)$ is the Hermite polynomial of order $m$ in $u$. It is important to realise that the functions $h_m(u)$ are orthonormal in the sense that

$$
\int_{-\infty}^{+\infty} h_m(u) h_n(u) \, du = \delta_{mn}.
$$

(33)

In these equations the symbols have their usual meanings. In particular, $w$ characterises the scale size of the beam at a plane, $R$ characterises the large-scale radius of curvature of the phase front, and $\theta$, the phase slippage between modes, characterises the form of the field as the beam propagates and diffracts. As has already been described in some detail [2], a mode set is not completely defined until the size $w_e$ and position of the waist are stated. Let us defer for a moment a discussion about how the waist size $w_e$ should be chosen. We do know, however, that at for a Gaussian-beam telescope the large-scale phase front of the field at the focal planes is flat, and therefore we can place the waists at these positions.

Having decided on the mode sets, we know that the scattering matrix of the whole system is just the product of the scattering matrices of the individual components. The components in this case are the two apertures and the free space paths between them; as usual the focusing effects of the ideal lenses are taken up by choosing the mode sets appropriately; that is to say the waists at all but the input plane are chosen according to the usual single-mode Gaussian-beam analysis. If we adopt this scheme there is no modal scattering associated with the ideal lenses.

The size of the waist at the input plane is still undetermined, and although any waist would produce a complete mode set in terms of which the field throughout the system could be represented, some particular input waist will be numerically more efficient than others. In a previous paper [2], we discussed at some length that the waist should be chosen to diagonalise as near as possible the scattering matrix, because in this case the Gauss-Hermite mode set chosen is a close as possible to the true eigenmodes of the system. In fact in that paper, we diagonalised the matrix to recover the true eigenmodes and eigenvalues which are known from classical analysis to be prolate spheroidal wavefunctions. Rather than choosing the mode set that best describes the natural modes of the optical system, we could choose the mode set that best represents the natural modes of the field. We know, however, that in the case of a perfectly incoherent source the two converge. In fact the way of choosing the waist described previously is based on the concept of incoherent modes, in the sense that we assumed that the intensity of the beam at a cross section is simply given by the incoherent sum of the individual mode intensities. It seems completely reasonable that the mode set for analysing the behaviour of a partially coherent field is precisely the same as that required for analysing a completely coherent field. We, therefore, assume that the mode set is also appropriate

456
for any partially-coherent field. We are of, course, merely talking about efficiency of convergence and so the precise choice is not critical anyway.

Hence, as before, we take the input waist of the optimum mode set to be

\[ w_1 = \left( \frac{\lambda f a_1}{\pi a_2} \right)^{1/2} \]  

(34)

where \( a_1 \) and \( a_2 \) are the radii of the apertures, and \( f \) is the focal length of the lenses. Moreover, the number of modes that should be used in the expansion is approximately the Fresnel number \( c \) where

\[ c = \frac{2 \pi a_1 a_2}{f \lambda} \]  

(35)

We are now in a position to derive the scattering matrix of the Gaussian-beam telescope. First we must calculate the scattering matrices of the apertures. By evaluating the field overlap integral over the output plane of each aperture [19, 20], and taking advantage of the fact that the large-scale phase front is flat at that point, we find that the scattering matrices are given by

\[ S_{m,n} = \int_{-\sqrt{c}k_t}^{+\sqrt{c}k_t} h_n(u) h_m(u) \, du \]  

(36)

where \( k_t = a/w \) is the normalised truncation. Because the mode set is, by definition, chosen so that the truncation at each stop is the same, we can easily write

\[ S_{m,n} = \int_{-\sqrt{c}}^{+\sqrt{c}} h_n(u) h_m(u) \, du \]  

(37)

Hence once we have chosen the Fresnel number, the scattering matrices of the two apertures are the same and given by the expression above. We could, of course, evaluate this matrix numerically, but we have found the following recurrence relationships useful. First we calculate the lowest-order coefficient through

\[ S_{00} = \text{erf}(\sqrt{c}) \]  

(38)

where \( \text{erf}(u) \) is the error function. Then, for all \( n + 1 \) odd we have

\[ S_{0,(n+1)} = S_{(n+1),0} = 0 \]  

(39)

and for all \( n + 1 \) even we have

\[ S_{0,(n+1)} = S_{(n+1),0} = -\left( \frac{2}{n+1} \right)^{1/2} h_0(\sqrt{c}) h_n(\sqrt{c}) \]  

(40)

Also, for \((m+1) + (n+1)\) odd

\[ S_{(m+1),(n+1)} = S_{(n+1),(m+1)} = 0 \]  

(41)

and for \((m+1) + (n+1)\) even

\[ \left( \frac{m+1}{n+1} \right)^{1/2} S_{m,n} - \left( \frac{2}{n+1} \right)^{1/2} h_{m+1}(\sqrt{c}) h_n(\sqrt{c}) \]  

(42)

These equations show that power does not scatter between odd and even ordered modes, as would be expected.

We also need the scattering matrices of the free-space paths. These are easily found because we know that the relationship between the focal planes of a Gaussian-beam telescope is a Fourier
Figure 2: The point-spread functions of Gaussian-beam telescopes having Fresnel numbers of 4, 8, and 16. The off-axis distance is normalised to the waist: $x/w$. Forty eight modes were used to construct these plots.

The transform or equivalently that the phase slippage is $\pi/2$ \cite{2}. The scattering parameters of the free-space paths therefore become

$$S_{m,n} = \left[ \cos \left( \frac{\pi m}{2} \right) + j \sin \left( \frac{\pi m}{2} \right) \right] \delta_{m,n}. \quad (43)$$

If we denote the scattering matrices of the free space paths by $S_f$ and the scattering matrices of the apertures by $S_a$, then the scattering matrix of the whole Gaussian beam telescope is simply given by the product

$$S = S_f S_a S_f S_a. \quad (44)$$

This matrix is extremely simple to generate using the above equations, and it is remarkable that it completely characterises the coherent and partially-coherent behaviour of the system. Also by raising the overall scattering matrix to some power we can calculate the effect of having a sequence of Gaussian-beam telescopes. This technique will be demonstrated shortly.

Now that we have generated the scattering matrix of a one-dimensional Gaussian-beam telescope, we can investigate its response to various different kinds of excitation. Before studying the propagation of partially-coherent radiation, it is worth while verifying the integrity of the scattering matrix by investigating the coherent behaviour. First of all we would like to plot the point-spread function.

The point-spread function is, of course, the output of the system when there is a delta function in the input plane. Moreover, for the Gaussian beam telescope, the input plane is at the position of a waist, where the phase front is flat. Calculating the mode coefficients in the one-dimensional case by evaluating the overlap integral, and then substituting the mode coefficients into the modal
expansion of the field, we find that the point-spread function, at the output plane, is given by

$$E_{p.s.f} (z, z') = \left( \frac{\sqrt{2}}{w} \right) \sum_m \sum_n S_{m,n} h_m \left( \frac{\sqrt{2}z'}{w} \right) h_m \left( \frac{\sqrt{2}z}{w} \right)$$

(45)

where $S_{m,n}$ are the elements of the scattering matrix and $z'$ is the position of the delta function in the input plane. In Fig. 2, we show the point-spread functions of a number of Gaussian-beam telescopes. This example demonstrates rather clearly how easy it is to calculate the point-spread function once the scattering matrix is known.

Rather than plotting the point-spread function for different input positions, it would be convenient to have some simple measure of its form. Classically, the crudest method use to use the Strehl ratio, where the Strehl ratio is defined as the height of the central peak normalised to the height of the peak at the central position. The argument being that because of the conservation of energy, any aberrations which spread the point-spread function will also reduce its height. In the case of Gaussian modes it is particularly easy to calculate the Strehl ratio. In fact it is given by

$$S = \frac{\sum_m \sum_n S_{m,n} h_m \left( \frac{\sqrt{2}z}{w} \right) h_m \left( \frac{-\sqrt{2}z}{w} \right)}{\sum_m \sum_n S_{m,n} h_m (0) h_m (0)}$$

(46)

Notice that the sign on one of the Hermite functions has changed to take into account the fact that, in our system, the peak in the point-spread function moves in the opposite direction to the position of the delta function. In Fig. 3 we show the Strehl ratio as a function of position for a number of different Fresnel numbers.

The Strehl ratio is independent of position over the whole of the field of view and this observation simply reflects the fact that we did not include any aberrations in our system. Some ringing can be seen at the edge of the input aperture, and this is a Gibbs phenomena due to the fact that we
are trying to describe the sharp edge of the input aperture with a finite number of modes. This ringing is not an optical effect, but exists merely because we are trying to represent the abrupt disappearance of the input field behind the input aperture. Moreover, it is only a second-order effect in the sense that it only represents a low-level ringing in the point-spread function. In fact Fourier theory gives us a limit to the percentage error in the height of the point-spread function as we move close to the edge of the field of view. The most impressive aspect of these plots is that they demonstrate that a small number of modes can represent the behaviour of the system over a large field of view, and this has significant implications for modelling the behaviour of imaging arrays.

Notice that in all of these plots, we used the expression

$$\frac{z}{w} = \frac{z}{a} \left( \frac{c}{2} \right)^{1/2}$$

in order to generate normalised scales.

We would now like to investigate the behaviour when a fully-incoherent source with a Gaussian intensity distribution is applied to the input. To perform this calculation, we require the coherence matrix. Using the one-dimensional form of the bimodal overlap integral and a cross-spectral density of the form

$$W(z', z) = f(z) \delta(z - z') = K^2 \exp \left( - \frac{2(z - z_0)^2}{\sigma^2} \right) \delta(z - z')$$

we find

$$C_{m,n} = K^2 \left( \frac{\sqrt{2}}{w} \right) \int_{-\infty}^{+\infty} \exp \left( - \frac{2(z - z_0)^2}{\sigma^2} \right) h_n \left( \frac{\sqrt{2}z}{w} \right) h_m \left( \frac{\sqrt{2}z}{w} \right) dz .$$

In Fig. 4, we show the intensity of the field at the output plane of a Gaussian-beam telescope when the Fresnel number is 4, 8, and 16. The normalised width of the effective input field distribution, $\sigma/w$, was taken to be 0.707. Superimposed on each plot is the cross-spectral density when the normalised reference position is 0.35: we could, trivially, have chosen any other reference position and the result would have been essentially the same.

For the purpose of generating a highly-incoherent input field distribution, we used 60 modes, but this large number of modes is not actually needed for the analysis. The main feature is, as would be expected, a slight spreading of the output intensity with decreasing Fresnel number and an increasing degree of coherence. What is not seen, because of normalisation, is the large amount of power lost, which would not be the case for a coherent field.

In addition to these plots it is also convenient to look at the behaviour when the input is a flat incoherent field of finite extent. In this case the elements of the coherence matrix are given by

$$C_{m,n} = \left( \frac{\sqrt{2}}{w} \right) \int_{-b}^{+b} h_n \left( \frac{\sqrt{2}x}{w} \right) h_m \left( \frac{\sqrt{2}x}{w} \right) dx ,$$

where $b$ is the extent of the field. Clearly, in the case where $b \to \infty$, the coherence matrix becomes diagonal, as expected. In Fig. 5 we show the intensity and cross-spectral density at the output plane when the Fresnel number is 4, 8 and 16.

Again the main feature of the plots is the smoothing of the highly-truncated input field and the increase in spatial coherence with decreasing Fresnel number. The cross-spectral densities should be compared with the point-spread functions shown in Fig. 2.

Finally, in Fig. 6 we show a sequence of plots where in each case the top-hat field distribution described above has been applied to a combination of Gaussian-beam telescopes all having a Fresnel number of 4. By simply raising the scattering matrix to the appropriate power, we show the effect of
Figure 4: In (a) we show the Gaussian intensity and cross-spectral density of the field at the input plane of a Gaussian-beam telescope. Sixty modes were used, for the purposes of the plot, to synthesize a nearly fully-incoherent field. The Gaussian input field has an effective normalised width of 0.707. In (b), (c) and (d) we show the intensity and cross-spectral density at the output plane when the Fresnel number is 4, 8, and 16.
Figure 5: In (a) we show the top-hat intensity and cross-spectral density at the input plane of a Gaussian-beam telescope. Sixty modes were used, for the purposes of the plot, to synthesize a nearly fully-incoherent field. The top-hat input field has an normalised half width of 1.414. In (b),(c) and (d) we show the intensity and cross-spectral density of the field at the output plane when the Fresnel number is 4, 8, and 16.
Figure 6: The intensity and cross-spectral density of the field at the output plane of a Gaussian-beam telescope when the Fresnel number is 4 and a fully-incoherent field with a top-hat intensity distribution is applied. (a), (b), (c) and (d) show the outputs when 1, 2, 4 and 8 identical telescopes are combined in series.

Having 2 telescopes, 4 telescope, and 8 telescopes in series. Clearly, this corresponds to adding more and more identical lens systems. It can be seen how coherence builds up due to mode filtering, and after 8 passes a fully-coherent Gaussian field is produced despite having started with fully-incoherent top-hat field. Indeed this is precisely the way in which coherence builds up in a laser cavity [21, 22, 23].

In order to demonstrate the behaviour of the system, we have used fully-coherent and fully-incoherent input fields. It is usual practice in classical optics to work in terms of Gauss-Schell sources [24, 25, 26], where both the intensity and the degree of coherence are Gaussian functions. It is then possible to change the degree of global coherence by varying the relative scale sizes. It would be straightforward to apply such a field to our system in order to investigate general behaviour, but it would not be particularly useful for practical applications. The examples we have studied here, are of course the two limiting extremes of this more general model. Despite these comments, it is interesting to note, that the natural modes of the Gauss-Schell source are Gauss-Hermite functions and the eigenvalues also have a particular form. It would be interesting to investigate the
implications of the scale sizes when such a beam is propagated through a Gaussian-beam telescope. In particular it should be possible to determine what size Gaussian intensity distribution would pass through a system with the minimum amount of loss. Moreover, such an analysis is closely related to the ability to use different approximations when calculating the scattering matrices of apertures. This issue has been discussed in some detail in the context of coherent fields [2], and it would be useful to extend the analysis to incoherent fields.

9 Conclusions

In this paper we have described in some detail a technique for calculating the behaviour of partially-coherent submillimetre-wave quasioptical systems. The technique, like its coherent equivalent, is based on the ability to scatter propagating modes at optical components. The field, instead of being described by a vector is described by a matrix, and this allows much more detail about a field to be traced through a system. In this way it is possible to distinguish between the natural modes of the field and the natural modes of the optical system through which the field is passing.

In the paper, we explained how to calculate the components of the coherence matrix from the known functional form of the cross-spectral density, and also how to propagate the coherence matrix in the forward and backward directions. We also gathered together a collection of tools for analysing performance when the scattering matrix is known. We illustrated the overall method by investigating the behaviour of one-dimensional Gaussian-beam telescopes. In a later paper we will use the same techniques to analyse in some detail the behaviour of arrays of multimode bolometers.

Not only is the method very powerful in a computational sense, but it also leads to considerable insight into the way multimoded partially-coherent quasioptical systems behave. We believe that the method can be extended to allow the noise performance of multibeam bolometer arrays to be calculated, to allow the experimental analysis of beams through maximum entropy techniques, and to allow the statistical properties of systems to be taken into account—for example to model the behaviour of a submillimetre-wave radio telescope with the atmosphere included.

References


Submitted to Infrared Physics and Technology.


