# Propagating Partially Coherent THz Fields Using Non-Orthogonal Over-Complete Basis Sets

R.H.Berry, S. Withington, M.P. Hobson, G. Yassin

Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge, UK.

April 30, 2003

**Abstract**

We discuss a general theoretical framework for representing and propagating coherent and partially coherent submillimetre-wave fields by means of mathematically over-complete sampled basis functions. One such basis set, Gabor modes, consists of overlapping translated and modulated Gaussian beams. In the paper, we present some extraordinary results from modal field reconstructions and propagations, including simple models of bolometer arrays. Our new formalism is currently being used to study other interesting schemes such as spatially compressed, and therefore non-orthogonal, Gaussian-Hermite polynomials.

## 1 Introduction

The THz region of the electromagnetic spectrum (100GHz-10THz, 3mm-30$\mu$m) offers numerous scientific and technical opportunities in areas as diverse as astronomy, atmospheric physics, medical imaging, gas phase spectroscopy, antiquities, and industrial process control. A major difficulty when designing instruments for these wavelengths is how to model, in an accurate and efficient manner, the behaviour of complicated coherent and partially coherent optical systems.

Two problems are encountered when attempting to model THz optical systems. First, the field of interest may be in a complicated state of coherence. For example, in the case of astronomical imaging arrays, the total field is comprised of the fields of the individual pixels, and depending on the type of system, these can be individually coherent or incoherent, and mutually coherent or incoherent. Second, when propagating and scattering a THz field, diffraction is important, and therefore ray tracing cannot be used, but conversely, full electromagnetic simulations are slow as multiple finely-sampled diffraction integrals are needed.

A partially coherent field may be regarded as a incoherent linear superposition of a set of fully coherent fields—the natural modes—and therefore the propagation and scattering of a partially coherent field can be regarded as the propagation and scattering of a number of coherent fields. Clearly, it will not be possible to analyse the behaviour of partially coherent systems unless the behaviour of fully coherent systems can be analysed efficiently. Gabor decompositions may offer a way out of this difficulty. It is known that Gabor decompositions can sample the real and Fourier spaces of space- and band-limited functions in a compact and elegant manner. In addition, a Gaussian field propagates in a modal fashion, and complex ray tracing can be used to trace the evolution of a Gaussian beam through a complex system of optical components. If Gabor decomposition can be extended to partially coherent fields, then partially coherent optical systems could be analysed easily, and many practical and theoretical developments would follow.

Previously we have analysed [1, 2, 3], partially coherent fields using a finite number of Gaussian-Hermite and Gaussian-Laguerre modes, basis sets in which the modes are mutually orthogonal but not complete, since an infinite number of such modes are needed to form a complete set. Such simulations have been very successful because the natural modes of a finite throughput Gaussian beam telescope are prolate spheroidal functions, closely related to such Gaussian based modes. In this paper we present a framework to quantify the degrees of freedom in a field a basis can support and we explicitly consider the numerical sampling of bases for optimal results. By these means we present an opportunity for the optimisation of sampling within modal analysis, ensuring that computational operations are reduced to a minimum.

We also demonstrate the feasibility of modelling the behaviour of partially coherent THz optical systems through the use of Gabor modes. In actual fact, the core of our analysis is not restricted to Gabor decompositions, but can be applied to any set of functions that form a non-orthogonal and *complete* or *over-complete* basis set. By complete we mean that the basis set supports all the degrees of freedom over a field's span without linear dependence between the basis functions, whereas for over-complete basis there are additional basis functions relative to the complete case and hence linear dependence in the basis set. In addition, we do not work with infinite-dimensional function spaces, but work with the practical reality of sampled, space-limited functions, which are spanned by finite dimensional vector spaces. Not only do we explore the behaviour of the basic mathematical structures involved, but we also provide an efficient numerical algorithm for decomposing partially coherent fields. The algorithm is demonstrated by propagating various fields through a simple two-dimensional Gaussian beam telescope. These models reproduce all of the well known classical behaviour, and achieve exceptional levels of numerical performance. Although in this paper we restrict ourselves to three-dimensional, scalar, paraxial propagation, the extension to vectorial paraxial propagation is straightforward essentially because paraxial vector fields separate in Cartesian coordinate systems. We also believe that it is straightforward to extend the basic technique to wide-angle, three-dimensional vector fields. Extended techniques to couple the power between fields and calculate the entropy associated with a field are presented in another of our forthcoming papers[4]. For the purpose of demonstrating the basic technique, we only propagate fields between Fourier planes, but in later work we will use complex ray tracing to calculate the field across intermediate surfaces.

## 2 Handling basis sets that are non-orthogonal and have varying degrees of completeness, including over-complete

We have previously shown[2] that partially coherent optical fields decomposed in orthogonal *complete* basis sets (modes) can be handled within the coherence matrix formalism. Now we shall present the analogous formalism for non-orthogonal *over-complete* basis sets. We decomposed optical fields into orthogonal complete bases by using continuous integral equations, but when such a method is implemented numerically the functions must be sampled at discrete points, and the sampling determines the number of degrees of freedom that must, potentially, be represented. The use of discrete representations is closely related to the way in which optical fields are measured experimentally. In order to deal with the degrees of freedom explicitly, therefore, we shall now work with discretely sampled fields. More specifically, we represent a one-dimensional optical field by a column vector $\boldsymbol{x}$ of length $N$. Throughout the paper, the notation $\boldsymbol{x}(i)$ is used to denote the $i^{th}$ element of the vector, $\boldsymbol{x}$. Consider an arbitrary over-complete and non-orthogonal basis set $\{\boldsymbol{e}_k\}$ where $k \in \{1, 2, ..., M\}$. In other words, we sample the field at $N$ points, and we wish to use $M$ basis functions to represent the field. Loosely speaking, over-determined implies $M > N$, uniquely-determined implies $M = N$, and under-determined $M < N$. That is to say, we certainly require enough degrees of freedom to describe all of the points in the original field. We wish to represent the field vector $\boldsymbol{x}$ in the form

$$\boldsymbol{x} = \sum_k a_k \boldsymbol{e}_k, \tag{1}$$

which is analogous to the continuous representation. We define the '*frame*' matrix, $\boldsymbol{S}$, as the sum of the matrices formed by taking the outer product of each basis function with its own conjugate transpose,

$$\boldsymbol{S} = \sum_k \boldsymbol{e}_k \boldsymbol{e}_k^{\dagger}, \tag{2}$$

where $(.)^\dagger$ indicates the conjugate transpose. At this point the physical and theoretical meaning of $\boldsymbol{S}$ and the associated operator are put aside, although it is beneficial to notice that for an orthonormal set $\boldsymbol{S} = \boldsymbol{I}$, indicating completeness, as would be true for an infinite set, $k \to \infty$, of discretely sampled Gaussian-Hermite modes. Equivalently, we can define a new $N \times M$ matrix $\boldsymbol{E}$ by collecting together the column vectors that represent the basis functions:

$$\boldsymbol{E} = [\boldsymbol{e}_1 \cdots \boldsymbol{e}_k \cdots \boldsymbol{e}_M]. \tag{3}$$

The frame matrix then simply becomes

$$\boldsymbol{S} = \boldsymbol{E}\boldsymbol{E}^\dagger. \tag{4}$$

We see that whereas the $N \times N$ matrix $\boldsymbol{S} = \boldsymbol{E}\boldsymbol{E}^\dagger$ is associated with completeness, the $M \times M$ matrix $\boldsymbol{R} = \boldsymbol{E}^\dagger \boldsymbol{E}$ is associated with orthogonality. Consider operating on $\boldsymbol{x}$ with $\boldsymbol{S}$ thus:

$$\boldsymbol{S}\boldsymbol{x} = \sum_k (\boldsymbol{e}_k^\dagger \boldsymbol{x}) \boldsymbol{e}_k, \tag{5}$$

where we use parentheses throughout the paper to emphasise when the enclosed quantity is an inner product, and therefore can be moved easily to a different position within an equation. Now assume that $\boldsymbol{S}$ is invertible in that there is some matrix $\boldsymbol{S}^{-1}$ such that,

$$\boldsymbol{S}^{-1}\boldsymbol{S} = \boldsymbol{I}. \tag{6}$$

The formal proof that $\boldsymbol{S}$ is invertible for a complete or over-complete set is presented in our theoretical paper [4]. It follows applying $\boldsymbol{S}^{-1}$ to (5) that

$$\boldsymbol{S}^{-1}\boldsymbol{S}\boldsymbol{x} = \boldsymbol{x} = \sum_k (\boldsymbol{e}_k^\dagger \boldsymbol{x}) \boldsymbol{S}^{-1} \boldsymbol{e}_k. \tag{7}$$

Now we identify a dual basis set $\{\tilde{\boldsymbol{e}}_k\}$. Each element is not a conventional dual vector, in the sense that because of over-completeness, the orthogonality condition,

$$\boldsymbol{e}_k^\dagger \tilde{\boldsymbol{e}}_l = \delta_{kl}, \tag{8}$$

is not satisfied. Rather, the dual basis functions are such that,

$$\tilde{\boldsymbol{e}}_k = \boldsymbol{S}^{-1} \boldsymbol{e}_k. \tag{9}$$

The dual function obtained through the use of $\boldsymbol{S}^{-1}$ is the canonical dual, the dual of lowest energy. Clearly in the case of a complete orthogonal set, $\boldsymbol{S} = \boldsymbol{I}$, and the dual vectors are the same as the basis vectors. According to (9), (7) becomes

$$\boldsymbol{x} = \sum_k (\boldsymbol{e}_k^\dagger \boldsymbol{x}) \tilde{\boldsymbol{e}}_k. \tag{10}$$

It is convenient to define a set of coefficients $b_k = \boldsymbol{e}_k^\dagger \boldsymbol{x}$, in which case the field expansion becomes

$$\boldsymbol{x} = \sum_k b_k \tilde{\boldsymbol{e}}_k. \tag{11}$$

We can now go through the same procedure again, but starting with the assumption that the inverse of $\boldsymbol{S}$ exists such that,

$$\boldsymbol{S}\boldsymbol{S}^{-1} = \boldsymbol{I}. \tag{12}$$

Hence,

$$\boldsymbol{S}\boldsymbol{S}^{-1}\boldsymbol{x} = \boldsymbol{x} = \sum_k \boldsymbol{e}_k \boldsymbol{e}_k^\dagger \boldsymbol{S}^{-1}\boldsymbol{x}. \tag{13}$$

It can be shown that because $\boldsymbol{S}$ is Hermitian, its inverse must be Hermitian:

$$\left(\boldsymbol{S}^{-1}\right)^\dagger = \boldsymbol{S}^{-1}, \tag{14}$$

and by using (9) we find

$$\tilde{\boldsymbol{e}}_k^\dagger = \boldsymbol{e}_k^\dagger \boldsymbol{S}^{-1}. \tag{15}$$

Substituting (15) into (13),

$$\boldsymbol{x} = \sum_k (\tilde{\boldsymbol{e}}_k^\dagger \boldsymbol{x}) \boldsymbol{e}_k. \tag{16}$$

Finally, we define the coefficients $a_k$ according to $a_k = \tilde{\boldsymbol{e}}_k^\dagger \boldsymbol{x}$ and obtain

$$\boldsymbol{x} = \sum_k a_k \boldsymbol{e}_k, \tag{17}$$

which on comparing with (1) is the form required. In summary, the set $\{\tilde{\boldsymbol{e}}_k\}$ forms an over-complete basis with the coefficients $b_k$, and the set $\{\boldsymbol{e}_k\}$ also forms an over-complete basis with coefficients $a_k$ and we have shown that a field can be decomposed into either of these bases.

# 3 Handling partially coherent fields

Using the expansions described in the previous section, we can now represent partially coherent fields using over-complete basis sets. Throughout this paper we assume, in accordance with the theory of analytic signals, that the complex sampled field components $\boldsymbol{x}(i)$ have small fractional bandwidth, $\delta f/f \ll 1$, and the spatial correlations of interest are characterised by second-order ensemble averages. Defining the space domain coherence matrix (SCM), $\boldsymbol{W}$, as the ensemble average, $\langle . \rangle$, of the outer product of the optical field vector $\boldsymbol{x}$ with its own conjugate:

$$\boldsymbol{W} = \left\langle \boldsymbol{x}\boldsymbol{x}^\dagger \right\rangle. \tag{18}$$

Each element of $\boldsymbol{W}$ is a measure of the complex spatial coherence between the field at two points. In this sense the SCM is the coherence matrix obtained when a set of delta functions, positioned at the sample points, is used as the basis set. The SCM contains sampled information that is the discrete analogy of the cross spectral density function. Using the results of the previous section, we can expand the SCM as follows. Substituting (11) and (17) into (18), we find

$$\begin{aligned} \boldsymbol{W} &= \left\langle \boldsymbol{x}\boldsymbol{x}^\dagger \right\rangle \\ &= \sum_k \sum_h \langle b_k a_h^* \rangle \tilde{\boldsymbol{e}}_k \boldsymbol{e}_h^\dagger \\ &= \sum_k \sum_h C_{kh} \tilde{\boldsymbol{e}}_k \boldsymbol{e}_h^\dagger, \end{aligned} \tag{19}$$

where the complex coefficients $C_{kh}$ constitute the coherence matrix elements in this representation. Substituting the definitions of the basis-function coefficients we find

$$\begin{aligned} C_{kh} &= \langle b_k a_h^* \rangle \\ &= \boldsymbol{e}_k^\dagger \boldsymbol{W} \tilde{\boldsymbol{e}}_h. \end{aligned} \tag{20}$$

The elements of the coherence matrix can therefore be calculated easily once the SCM, and the dual basis vectors are known. Here we have expressed the coherence matrix in terms of the correlations between the coefficients of the Gabor functions and the coefficients of the associated dual basis. Later we will discuss the alternative representation using correlations between the Gabor functions and themselves. First we will examine why the representation in (20) has a number of attractions. It can be seen from (20) that the overall structure reduces to simple and well-known forms in all of the various limits that can be imagined. For example, suppose that the source comprises a fully-incoherent surface with uniform intensity. In this case $\boldsymbol{W} = \boldsymbol{I}$, and according to (20) the coherence matrix elements become

$$C_{kh} = \boldsymbol{e}_k^\dagger \tilde{\boldsymbol{e}}_h. \tag{21}$$

For an *over-complete* set, the basis functions and the associated dual functions are not orthogonal; however for a *complete* non-orthogonal set, the basis functions and the dual functions are mutually orthogonal, and in this case,

$$C_{kh} = \delta_{kh}, \tag{22}$$

and the conventional dual condition (8) holds. (22) is important as it shows that, for a non-orthogonal and complete set, a completely spatially incoherent source excites all basis functions incoherently and equally, as is the case for orthonormal sets. In the case of an over-complete basis set, a fully-incoherent source leads to correlations between the basis vectors, due to their linear dependence, as would be expected on physical grounds.

The expansion described by (19) has many convenient qualities, but it is inconvenient from an optical point of view, as we would like to propagate partially coherent fields by propagating the basis functions only. If we use (19) it would be necessary to propagate the dual functions as well as the basis functions. Suppose, therefore, that we use the form

$$\begin{aligned} \boldsymbol{W} &= \langle \boldsymbol{x}\boldsymbol{x}^\dagger \rangle \qquad (23) \\ &= \sum_k \sum_h \langle a_k a_h^* \rangle \, \boldsymbol{e}_k \boldsymbol{e}_h^\dagger \\ &= \sum_k \sum_h C_{kh} \boldsymbol{e}_k \boldsymbol{e}_h^\dagger, \end{aligned}$$

where the matrix coefficients are given by

$$C_{kh} = \tilde{\boldsymbol{e}}_k^\dagger \boldsymbol{W} \tilde{\boldsymbol{e}}_h. \tag{24}$$

This decomposition has the advantage that it is only necessary to propagate basis functions, but has the disadvantage that even for non-orthogonal complete sets, the coherence matrices do not take on simple forms for fully incoherent fields. It does, however, lead to an insight into the meaning of the frame matrix $\boldsymbol{S}$. Suppose that we excite all of the basis functions incoherently and equally, which we cannot do in practice for an over-complete set because of linear dependence. In this case, the coherence matrix becomes $\boldsymbol{C} = \boldsymbol{I}$, which can be substituted into (23) to yield the implied spatial correlations, this substitution gives,

$$\begin{aligned} \boldsymbol{W} &= \sum_k \sum_h \delta_{kh} \boldsymbol{e}_k \boldsymbol{e}_h^\dagger \qquad (25) \\ &= \sum_k \boldsymbol{e}_k \boldsymbol{e}_k^\dagger, \end{aligned}$$

which is the frame matrix $\boldsymbol{S}$. This result is not surprising when one remembers that $\boldsymbol{e}_k \boldsymbol{e}_k^\dagger$ gives the spatial correlations for each basis vector, and if two mutually incoherent fields are added, regardless of the state of coherence of each, the overall coherence matrix is simply the sum of the two individual coherence matrices. Hence, the frame matrix is simply the coherence matrix of a sum of individually fully coherent, but mutually incoherent basis fields. In short, the frame matrix contains information about the spatial correlations, due to linear dependence, associated with all of the basis vectors.

The explicit use of the SCM to calculate the coherence matrix elements provides a method for representing coherent, fully incoherent and partially coherent fields using over-complete bases. A major advantage of the scheme is that SCM can be calculated for a variety of different fields, and complicated forms assembled. For a fully coherent field described by vector $\boldsymbol{x}$, the coherence matrix is trivially $\boldsymbol{W} = \boldsymbol{x}\boldsymbol{x}^{\dagger}$. For a fully incoherent uniform field $\boldsymbol{W} = \boldsymbol{I}$, whereas for a fully incoherent field, with some spatially varying intensity distribution, the SCM is simply diagonal, with the diagonal elements scaled according to the intensity.

By noticing that we can add mutually incoherent fields by adding coherence matrices, we can set up the fields associated with numerous practical situations. For example, we can model a noisy aperture by adding to the scattered field, the field associated with the blackbody radiation of the aperture.

# 4 Properties of the frame matrix

The invertibility of $\boldsymbol{S}$ to form the $\boldsymbol{S}^{-1}$ matrix is vital to the existence and formation of the dual frame. It can be demonstrated using elegant linear algebra that $\boldsymbol{S}$ is invertible for critically or over sampled frames [4]. The invertibility of $\boldsymbol{S}$ appears in a different form in Daubechies'[5] definition of an associated frame operator for continuous frames and use of Bessel's inequality. Here we give explicit consideration to the singular value decomposition (SVD) of the frame matrix leading us to the same conclusion. We shall now consider two hermitian matrices formed from matrix $\boldsymbol{E}(3)$,

$$\boldsymbol{S} = \boldsymbol{E}\boldsymbol{E}^{\dagger}, \tag{26}$$

and

$$\boldsymbol{R} = \boldsymbol{E}^{\dagger}\boldsymbol{E}, \tag{27}$$

Several important properties of the basis can be investigated by considering the SVD of these matrices and the matrix, $\boldsymbol{E}$. $\boldsymbol{E}$ can be expressed in the form of a singular value decomposition as described in numerical recipes[6] such that

$$\boldsymbol{E} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\dagger}, \tag{28}$$

where both matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary, $\boldsymbol{U}$ having dimensions $N \times N$ and $\boldsymbol{V}$ having dimensions $M \times M$ and $\boldsymbol{\Sigma}$ is a matrix $N \times M$. The columns of $\boldsymbol{U}$ contains the left singular values of $\boldsymbol{E}$, the vectors that span the range of $\boldsymbol{E}$. $\boldsymbol{V}$ contains the right singular vectors of $\boldsymbol{E}$. $\boldsymbol{\Sigma}$ is diagonal with the $i^{th}$ diagonal element, $\sigma_i$. The elements of $\boldsymbol{\Sigma}$ are ordered in descending size and the first $min(M, N)$ diagonal elements are non-zero with the $i^{th}$ element corresponding to the $i^{th}$ column of $\boldsymbol{U}$, $\boldsymbol{u}_i$, is the corresponding vector that spans the range of $\boldsymbol{E}$. The condition of the vector space spanned by $\boldsymbol{E}$ is defined as the ratio of the smallest and largest values, we have considered an ill conditioned situation to arise when this ratio $< 10^{-5}$. The values of $\boldsymbol{\Sigma}$ where $\sigma_i$ is zero (or as good as numerically) correspond to the $i^{th}$ column of $\boldsymbol{V}$, $\boldsymbol{v}_i$, the column of $\boldsymbol{V}$ that gives a vector that span the null space of $\boldsymbol{E}$.

Using the unitary properties of $\boldsymbol{U}$ and $\boldsymbol{V}$ it follows that,

$$\begin{aligned} \boldsymbol{S} &= \boldsymbol{E}\boldsymbol{E}^{\dagger} \\ &= \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\dagger}\boldsymbol{V}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{U}^{\dagger} \\ &= \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{U}^{\dagger}, \end{aligned} \tag{29}$$

and

$$\begin{aligned} \boldsymbol{R} &= \boldsymbol{E}^{\dagger}\boldsymbol{E} \\ &= \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{U}^{\dagger}\boldsymbol{U}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{V}^{\dagger} \\ &= \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{V}^{\dagger}. \end{aligned} \tag{30}$$

Given $\boldsymbol{\Sigma}$ is diagonal the first $min(M, N)$ diagonal elements of both $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}$ and $\boldsymbol{\Sigma}^{\dagger}\boldsymbol{\Sigma}$ are identical with an $i^{th}$ diagonal element of $\sigma_i^2$. So the smaller matrix of $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}$ and $\boldsymbol{\Sigma}^{\dagger}\boldsymbol{\Sigma}$ has eigenvalues satisfying $\sigma_i^2$ and the larger matrix has the same first $min(M, N)$ eigenvalues with the rest numerically equivalent to zero.

From equation(29)

$$\begin{aligned} \boldsymbol{U}^{-1}\boldsymbol{S}\boldsymbol{U} &= \boldsymbol{U}^{-1}\boldsymbol{E}\boldsymbol{E}^{\dagger}\boldsymbol{U} \\ &= \boldsymbol{U}^{-1}\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{U}^{\dagger}\boldsymbol{U} \\ &= \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}, \end{aligned} \tag{31}$$

and likewise,

$$\begin{aligned} \boldsymbol{V}^{-1}\boldsymbol{R}\boldsymbol{V} &= \boldsymbol{V}^{-1}\boldsymbol{E}^{\dagger}\boldsymbol{E}\boldsymbol{V} \\ &= \boldsymbol{V}^{-1}\boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{U}^{\dagger}\boldsymbol{U}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{V}^{\dagger}\boldsymbol{V} \\ &= \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}. \end{aligned} \tag{32}$$

Since both $\boldsymbol{S}$ and $\boldsymbol{R}$ are hermitian the columns of $\boldsymbol{V}$ are the normalised eigenvectors of $\boldsymbol{R}$, $\boldsymbol{v}_i$, $i = 1...M$; and the columns of $\boldsymbol{U}$ are the normalised eigenvectors of $\boldsymbol{S}$, $\boldsymbol{u}_i$, $i = 1...N$. The matrix $\boldsymbol{S}$ measures the completeness of the basis set, whereby if the number of non-zero eigenvalues of $\boldsymbol{S}$ equals $N$ then the basis is complete or over-complete and can represent all the degrees of freedom in the signal. The matrix $\boldsymbol{R}$ measures the orthogonality of the basis set.

# 5 Gabor Basis Sets

An over-complete non-orthogonal basis can be formed by a subset of frames known as the Gabor frames. Indeed much of the development of the formulation to handle frames was done with consideration to this particular set of bases. In Gabor bases the basis set modes are generated from a single mother function or *atom*$(\boldsymbol{g})$, translated and modulated. Now we shall describe the modulation operator and translation operator which together generate the frame from the mother atom. We are again considering the 1-D field sampled by a vector of length $N$, our mother atom and consequently basis modes are therefore also described by vectors of length $N$ . To form a frame, a translation spacing$(a)$ and a modulation spacing$(b)$ are chosen. To avoid boundary artifacts arising from assumed periodicity in the field signal and basis both $a$ and $b$ are chosen to be integer factors of $N$. Two other integers representing the number of modes arising from the operators on field vector are defined, $\tilde{a} = \frac{N}{a}$ and $\tilde{b} = \frac{N}{b}$. The expression $(p \mod q)$ represents the modulo operation associated with integer division, giving the integer remainder when $p$ is divided by $q$. The translation operator $T$ is,

$$\boldsymbol{T}_{va}\boldsymbol{g}(j) = \boldsymbol{g}((j - va) \mod N) \qquad v \in Z, \tag{33}$$

and the modulation function $M$ operates such that

$$\boldsymbol{M}_{vb}\boldsymbol{g}(j) = e^{-\frac{2\pi i.j\,((v+v_0) \mod N))b}{N}}\boldsymbol{g}(j) \qquad v \in Z, \tag{34}$$

where $v_0$ is an phase offset introduced to ensure that the phase differences between modes are zero on the input plane for at a defined optical axis, this method for assigning an optical axis can only be used when the optical axis is chosen to lie on a sampling point and not between them. To define an optical axis at a half-integer point we use $(v_0 = 0)$ and multiply every element in $\boldsymbol{g}$ by a complex phase factor, this is slightly more costly computationally so to maximise the size of the largest system that can be realistically examined it is advisable to chose the sampling such that the optical axis lies at a sampling point, giving an integer value of $v_0$. Using the Gabor basis functions the frame matrix elements can again be calculated as the outer product of the basis functions as in equation(4) giving,

$$\boldsymbol{S}_{jk} = \sum_{n=0}^{\tilde{a}-1}\sum_{m=0}^{\tilde{b}-1} \boldsymbol{M}_{mb}\boldsymbol{T}_{na}\boldsymbol{g}(j)(\boldsymbol{M}_{mb}\boldsymbol{T}_{na}\boldsymbol{g}(k))^{\dagger} \tag{35}$$

As yet no information about the functional form $\boldsymbol{g}$ is assumed other than it must have suitable support for a complete basis to be formed. Given our definitions above of $\boldsymbol{T}_{na}$ and $\boldsymbol{M}_{mb}$, for Gabor frames, $\boldsymbol{S}$ will clearly have a sparse banded form, this structure can be exploited as many optimised matrix manipulation algorithms exist for such matrices. Additionally and importantly the Gabor frame operator, associated with the frame matrix, commutes with both the modulation and the translation operator such that when $\boldsymbol{S}^{-1}$ exists (i.e. for complete and over-complete frames) we see,

$$\begin{aligned}\tilde{\boldsymbol{g}}_{mn} &= \boldsymbol{S}^{-1}\boldsymbol{g}_{mn} \\ &= \boldsymbol{S}^{-1}\boldsymbol{T}_{na}\boldsymbol{M}_{mb}\boldsymbol{g} \\ &= \boldsymbol{T}_{na}\boldsymbol{M}_{mb}\boldsymbol{S}^{-1}\boldsymbol{g} \\ &= \boldsymbol{T}_{na}\boldsymbol{M}_{mb}\tilde{\boldsymbol{g}}.\end{aligned} \tag{36}$$

Consequently only the mother dual associated with the mother function need be calculated and the entire basis set can be generated using the same operators. Gabor originally presented a Gaussian mother function atom as suitable with the form,

$$\boldsymbol{g}(x) = A_0 exp(-\pi\frac{x^2}{2a^2}), \tag{37}$$

where $A_0$ is a scaling factor and $a$ is again the translation spacing of the basis as described above. Gabor modes have been used for coherent signal coefficient decomposition and image analysis and such applications and implementations of frames are discussed in papers such as Feichtinger's papers[7, 8].

# 6 Implementation

The inversion of the frame matrix was not performed directly but a conjugate gradient method (CGM), as described in Numerical Recipes [6], was used to minimise the residue,

$$\epsilon = \left\|\boldsymbol{S}\tilde{\boldsymbol{g}} - \boldsymbol{g}\right\|, \tag{38}$$

and hence determine the mother dual for a 1-D system. The choice of the norm operator $\|.\|$ in (38) determines the form of the dual and whether null space vectors of $\boldsymbol{S}$ are added to the canonical dual. This method avoids any direct intermediate calculation of the inverse of $\boldsymbol{S}$ and was found to converge very rapidly. The success of this process of minimisation is directly related to the results of the SVD of $\boldsymbol{E}$ and the invertibility of $\boldsymbol{S}$. Additionally this method was able to find an approximate solution when $\boldsymbol{S}$ is not invertible, i.e. for under-complete frames, when an approximate form for the dual is found through the use of the analogous form of (36). Having obtained the mother dual and consequently the dual basis set, successful decompositions and reconstructions of fields were performed using the coherence matrix form given in (23). This is the form of the coherence matrix which contains the coherence elements associated with the Gabor basis functions themselves and was used because, as described next, these basis functions were then propagated.

## 6.1 Propagating the Gabor basis functions

The Gabor basis functions are Gaussian in intensity and of particular interest as many techniques for propagating Gaussian beams exist in this field. Propagation is especially simple for paraxial quasi-optical systems. Complex ray tracing of Gaussian beams is also extremely useful. An important subset of frames is that of critically sampled frame where $N = M$ (and hence $N = ab$) and additionally the Gabor mother function takes the form of (37). An oversampled frame has $N < M$ (and hence $N > ab$). The mathematical continuous form of a critically sampled frame was shown to be complete in Gabor's original paper of 1946 proposing such a basis set [9], however the duals were not deduced until 1980 by Baastians [10]. Propagation between Fourier planes is trivial for these frames if condition (37) is satisfied because the frame at the input plane propagates to another Gabor frame at the Fourier plane formed at the focal plane of a lens. The propagation was perform by coherence matrix element swapping, as each modulation function in the original basis set propagates to a translation function in the Fourier frame and vice-versa. This method of Fourier transform elegantly leads to the same information sampling across both the input and transform plane and is extremely fast. Fourier analysis can of course only be applied to paraxial quasi-optical systems, however such systems are of considerable interest and worth considering. We analysed several 1-dimensional systems using such frames including systems where the Fourier plane was limited. Additionally we then analysed some 2-dimensional systems through the use of 1-dimensional operations described, the justification and method to do this is now given.

### 6.2 Two Dimensional Fields

An important subset of two dimensional fields $\boldsymbol{F}(x, y)$ is where the field is separable in that the two components, $\boldsymbol{X}(x)$ and $\boldsymbol{Y}(y)$ are independent. Such a two dimensional field is simply the product of one dimensional fields, $\boldsymbol{F}(x, y) = \boldsymbol{X}(x)\boldsymbol{Y}(y)$, and can naturally be analysed using the product of two one dimensional basis function sets. Any components and the propagation process itself must not affect the separability of the field. Non-separable basis functions can be analysed within the mathematical framework; however, the numerical cost involved in any practical implementation calculating the coherence matrices in such a basis set is high and for a realistic number of sampling points often prohibitively so. Considering a two dimensional system as the product of two one dimensional systems creates a problem of the order $2 \times N^2$ whereas to consider the initial field as non-separable is a problem of the order $N^4$, when double inner products are necessary within the analysis the high cost of this is very apparent. Components such as circular apertures and Zernike phase screens are non-separable in such an analysis. Fortunately we can analyse such components within the separable framework through explicit consideration of the discrete sampling used. For any field sampled in Cartesian coordinates (pixellated) a non-separable field can be expressed as a sum of separable sampled fields. This trivially follows from the pixel basis being the outer product of the $x$ and $y$ co-ordinate systems. Likewise, discrete sampling of components such as circular apertures allows them to be analysed in the same way. As a pixellated system tends from separable into two 1-D functions to only separable on a pixel scale the cost of the computation tends to the higher $N^4$ limit. Each of the separable fields within the total field can be analysed separately and the coherence matrix elements of constituent fields summed at the appropriate output plane. To determine the separable form of any continuous field we can consider another SVD performed on the matrix $\boldsymbol{F}$ where the elements of $\boldsymbol{F}$ are $F_{ij} = \boldsymbol{F}(x_i, y_j)$. If the SVD is performed, using the same notation as before for the SVD, such that $\boldsymbol{F} = \boldsymbol{U\Sigma V}^\dagger$, then the pixellated form of the function sampled by an $H \times H$ matrix is given by,

$$\boldsymbol{F}(x, y) = \sum_{n=1}^{H} \sigma_n \boldsymbol{u}_n(x) \boldsymbol{v}_n(y)^\dagger. \tag{39}$$

For many fields the field or a good approximation to it can be given by,

$$\boldsymbol{F}(x, y) = \sum_{n=1}^{J} \sigma_n \boldsymbol{u}_n(x) \boldsymbol{v}_n(y)^\dagger, \tag{40}$$

where $J < H$, reducing the complexity and increasing the computational limits of the problem. Some simple 2-dimensional fields formed as the outer product of two 1-dimensional fields were investigated.

## 7 Results

### 7.1 The form of the dual

The form of some of the dual mother atoms for a fixed value of $N$ but with varying values of $a$ and $b$, forming critically or over-complete frames, are shown in Fig.1. The form of the duals can be seen to differ dramatically. Many of the duals can be seen to contain discontinuities especially in critically sampled frames and careful consideration of these is necessary within numerical simulations. Additionally it can be seen that increasingly over-sampled systems have smoother dual functions.

### 7.2 Investigating the completeness of the basis set

The eigenvalues of $\boldsymbol{\Sigma\Sigma}^\dagger$ and $\boldsymbol{\Sigma}^\dagger\boldsymbol{\Sigma}$ were found using SVD and the highest and lowest eigenvalues of both $\boldsymbol{S}$ and $\boldsymbol{R}$ plotted for a variety of frames with a fixed number of samples in a field. The results for two frames

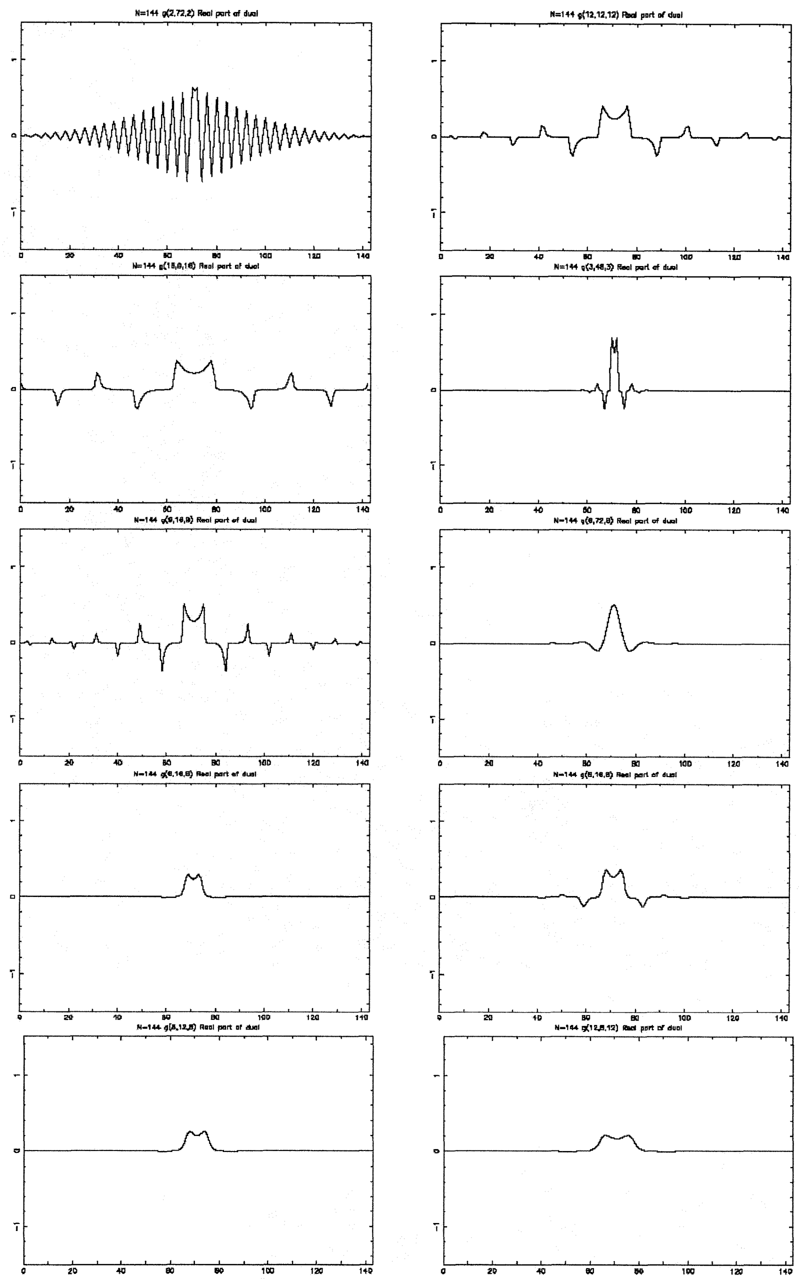

Figure 1: From top to bottom and left to right the frames are defined with a mother atom of the form given in 37 and the values of (N, a, b) are given to define the frames as (144,2,72), (144,12,12), (144,9,16), (144,6,72), (144,16,9), (144,3,48), (144,6,16), (144, 8, 16), (144,8,12), (144,12,8)

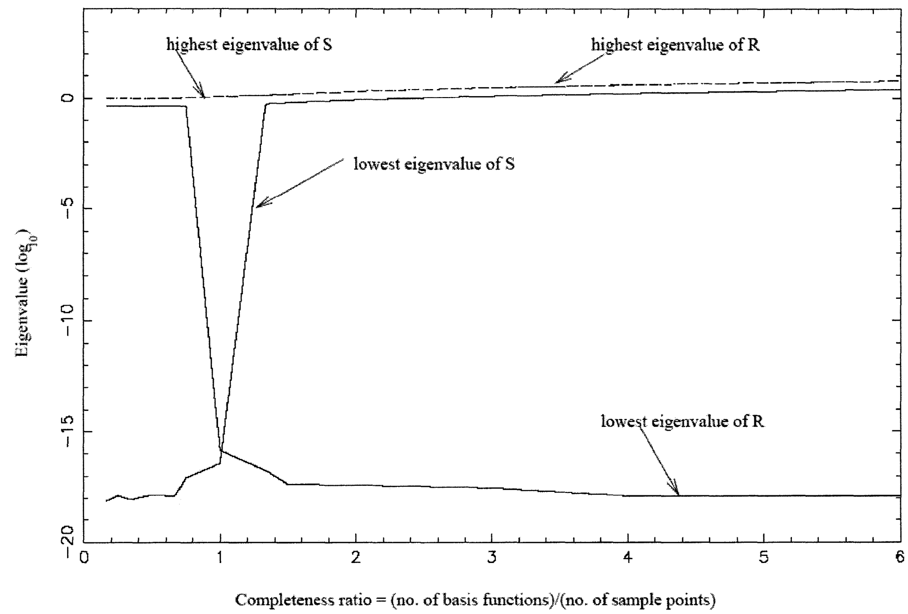


Figure 2: The highest and lowest eigenvalues of both $\boldsymbol{S}$ and $\boldsymbol{R}$. $N = 144$, $a = 12$, $b \in \{2, 3, 4, 6, 8, 9, 12, 16, 18, 36, 72\}$ waist satisfying (37). The critically sampled frame is not complete.

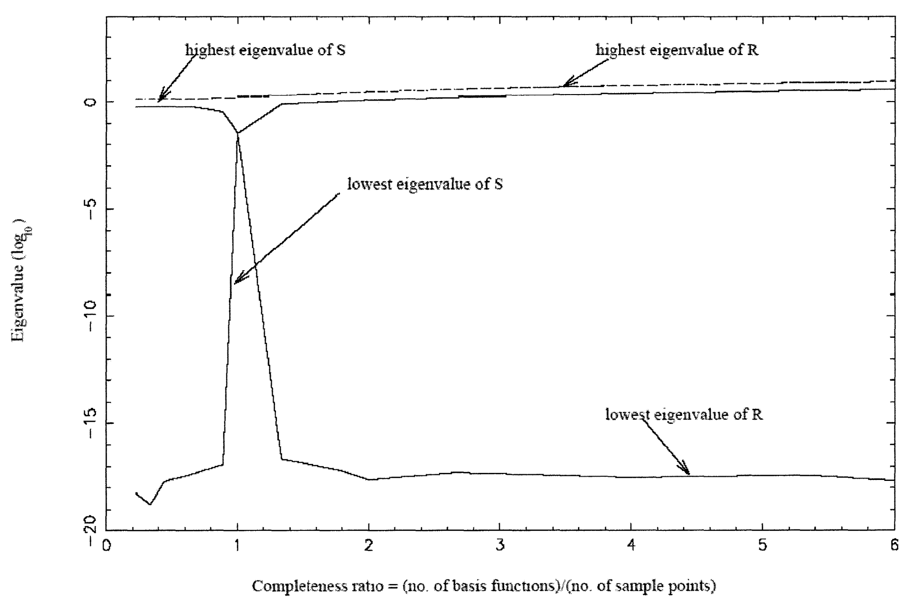


Figure 3: The highest and lowest eigenvalues of both $\boldsymbol{S}$ and $\boldsymbol{R}$. $N = 144$, $a = 9$, $b \in \{2, 3, 4, 6, 8, 9, 12, 16, 18, 36, 72\}$ waist satisfying (37). The critically sampled frame is complete.

with $N = 144$ are considered here. The first frame with $a = 12$ and $b \in \{2,3,4,6,8,9,12,16,18,36,72\}$ is shown in Fig.2, a second frame with $a = 9$ and $b \in \{2,3,4,6,8,9,12,16,18,36,72\}$ is shown in Fig.3. The waist for every frame was taken to satisfy (37). In these investigations we fixed the number of translational modes and changed the completeness of the frame by varying the number of modulations used.

## 7.3 Quantifying the degree of completeness of a frame

The information that can be supported was invested for under determined, critically sampled and over determined frames. The number of samples of the basis functions was fixed. The ratio of the number of non-singular eigenvalues from the SVD to samples of the basis function was plotted against the translation spacing $a$ for various values of modulation spacing $b$. By the assumed periodicity of our method we are restricted to certain values of $a$ and $b$, however other frames with different numbers of samples can provide other values. Data for the parameters $N = 144$, $a \in \{2,3,4,6,8,9,12,16,18,36,72\}$ and $b \in \{6,8,9,12,16,18,36,72\}$ was calculated, the waist for every frame was taken to again satisfy (37). The results are shown in Fig.(4).

## 7.4 Analysing the results investigating the completeness of the frame

An interesting feature arises in our results; when investigating critically sampled Gabor frames we expect $\boldsymbol{S}$ to be invertible and to support N degrees of freedom. However the results of Fig.2 show that the SVD results in a zero lowest eigenvalue, on closer investigation it was found that only this eigenvalue was zero and thus a single degree of freedom is not represented. However the frame investigated in Fig.3 was found to be complete and support all degrees of freedom in the signal. The continuous form of the traditional Gabor mother atom has been shown to form a frame that spans the space and the discontinuities seen in the duals of critically sampled frames in Fig.1 can be described mathematically in a continuous analysis[10]. However this missing degree of freedom anomaly arises from a discretely sampled implementation of such discontinuous duals. Critically sampled frames can be said to be *unstable* due to this behaviour. Clearly as critical sampling is approached there is a change in behaviour of the frames in Fig.2 and Fig.3, where by the lowest eigenvalue dips relative to the trend. It is simply the size of this effect that determines whether $\boldsymbol{S}$ is invertible and the frame spans the space. For our paraxial propagation investigations we investigated critically sampled frames, ensuring that the frames were complete or that there was no power in any missing degree of freedom. Ideally in the future we wish to quantify the conditions under which critically sampled frames are also complete frames. However it should be emphasised that this is a minor consideration as we have found and used complete, critically sampled frames with great success.

## 7.5 A 1-D illustration of the Van-Cittert Zernike Theorem

Fig.5 shows the simulations of two systems. The top left plot shows the reconstruction of two fully self and mutually coherent Gaussians, the intensity is shown by the solid line and lies concurrently with a dashed line showing the coherence function. The reconstruction of the field of length $N = 81$ was performed in a Gabor basis set with $a = 9$ and $b = 9$, i.e. using a basis set consisting of Gaussians of a different size to those in the field. The field was then propagated to the Fourier plane and the results of the reconstruction are shown in the top right plot. The bottom left plot shows the reconstruction of two fully self and mutually incoherent Gaussians, and two dashed lines are plotted plotting the coherence function relative to the centre of each Gaussian. The intensities of the Gaussians were identical to the coherent case as was the basis set used for the decomposition. This field was propagated to the Fourier plane and the results are shown in the bottom right plot. It should be emphasised that the input fields shown are not the simply the inputed field but the basis function coherence matrix decomposition and reconstruction of this field.

There are several outstandingly successful features to notice in these results. The decompositions successfully retain all the degrees of freedom in the field, including the representation of the delta functions associated with the coherence function of the incoherent Gaussians. The coherent Fourier plane results show the fringes as expected from regular Fourier theory. Although the fringes in Fig.5 are extremely sharp, they are *not* under sampled but simply only contain the original information from the field, without artificial smoothing from oversampling. By only representing the degrees of freedom in the original field

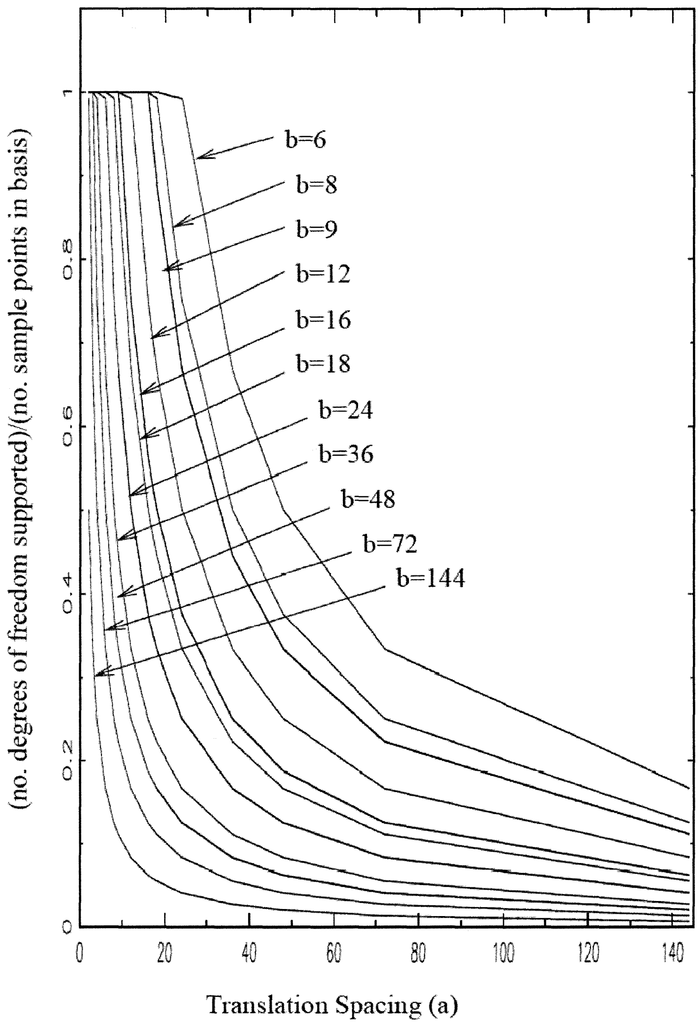


Figure 4: The degrees of freedom a frame with the parameters $N = 144$, $a \in \{2, 3, 4, 6, 8, 9, 12, 16, 18, 36, 72\}$ and $b \in \{6, 8, 9, 12, 16, 18, 36, 72\}$ is shown, the waist for every frame was taken to again satisfy (37). An important featur of this plot is the region where the plots tend to a value of 1 on the vertical axis; it can be seen some critically sampled frames are complete whereas others have a single missing degree of freedom.

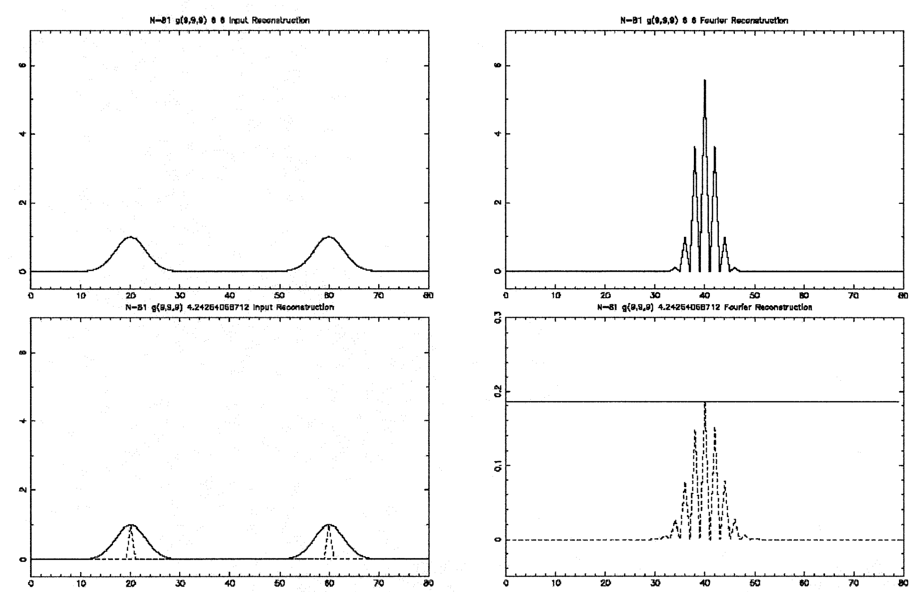


Figure 5: Top Left: Two coherent Gaussian sources, reconstructed in a basis set of Gabor functions consisting of Gaussians of a *different* width to the sources, at input plane. Top Right: The coherent Gaussian sources at the image (Fourier) plane. Bottom Left: Two mutually and self incoherent Gaussian sources, reconstructed in a basis set of Gabor functions consisting of Gaussians of a *different* width to the sources, at the input plane. Bottom Right: The incoherent Gaussian sources at the image (Fourier) plane. The solid line in all plots shows the intensity of the function and the dotted line the coherence function. At the input plane for the coherent Gaussians the coherence function and the intensity line are concurrent. At the input plane, for the incoherent sources, two coherence functions are plotted relative to the centre of each of the Gaussians.

we are not supporting redundant information and are dramatically increasing the power and efficiency of the simulations. There are equally interesting results from the incoherent system's Fourier plane, as expected from a fully incoherent field the Fourier intensity is uniform and contains no spatial features. Additionally the form of the fringes in the Fourier coherence function can be seen to be the same as those in the coherent Fourier intensity; illustrating the well known results from the Van-Cittert Zernike theorem whereby the forms of the coherence and the intensity are swapped at the Fourier plane.

### 7.6 Diffraction Limited 1-D System

A simple finite throughput quasi-optical system was simulated. A fully coherent top hat function was decomposed and propagated to the Fourier plane, where an aperture was placed to limit the throughput. The limited field was then propagated to the image plane and reconstructed. The results for varying size of apertures are shown in Fig.6. It should be noted that all the fields shown have resulted from reconstructions, where the coefficients associated with the basis Gabor functions have been calculated and the basis functions propagated. The top left hand plot is the reconstruction where an infinite aperture has been used at the Fourier plane and hence the original field has been reconstructed at the subsequent image plane.

Again all degrees of freedom in the original field have been preserved in the decomposition and remarkably the stepped edges of the top hat function have been perfectly reproduced. Even using a large number of Gaussian-Hermites the edges are not reproduced and Gibbs oscillations are seen.

### 7.7 2-D separable systems, including simple bolometer and horn array models

We modelled a simple bolometer system and transformed the field to the Fourier plane. The intensity and coherence function of the original field represented in a frame basis and the corresponding results in the Fourier plane are shown in Fig.7. A simple horn array was then modelled with the same intensity function as the bolometer array modelled in Fig.7. The horn array results are shown in Fig.8.

## 8 Conclusions

We conclude that the mathematical framework, presented in this paper, provides a method for:

- Explicitly handling sampling within modal decompositions ensuring no redundancy in the field information is propagated. By optimising the sampling to the field we avoid unnecessary calculations that are computationally expensive while still maintaining the accuracy of our simulations.
- Simulating fully coherent, fully incoherent and partially coherent field
- Elegantly handling new basis sets, previously avoided because of the difficulties handling non-orthogonal over-complete basis sets.
- Quantifying the range of basis sets and quantifying improvements to basis sets by measuring ratio of the degrees of freedom in the field to those supported by the basis.

The simulations presented demonstrate the success of this framework by:

- Producing field reconstructions which retain all the degrees of freedom present in the initial field.
- Reconstructing delta functions and sharp edges in fields, field features that a finite set of Gaussian-Hermite basis modes can not support.

Additionally our work in this area has derived relations for the power coupled between two partially coherent fields and the entropy of a partially coherent field within the framework of over-complete basis sets. To be published [4].

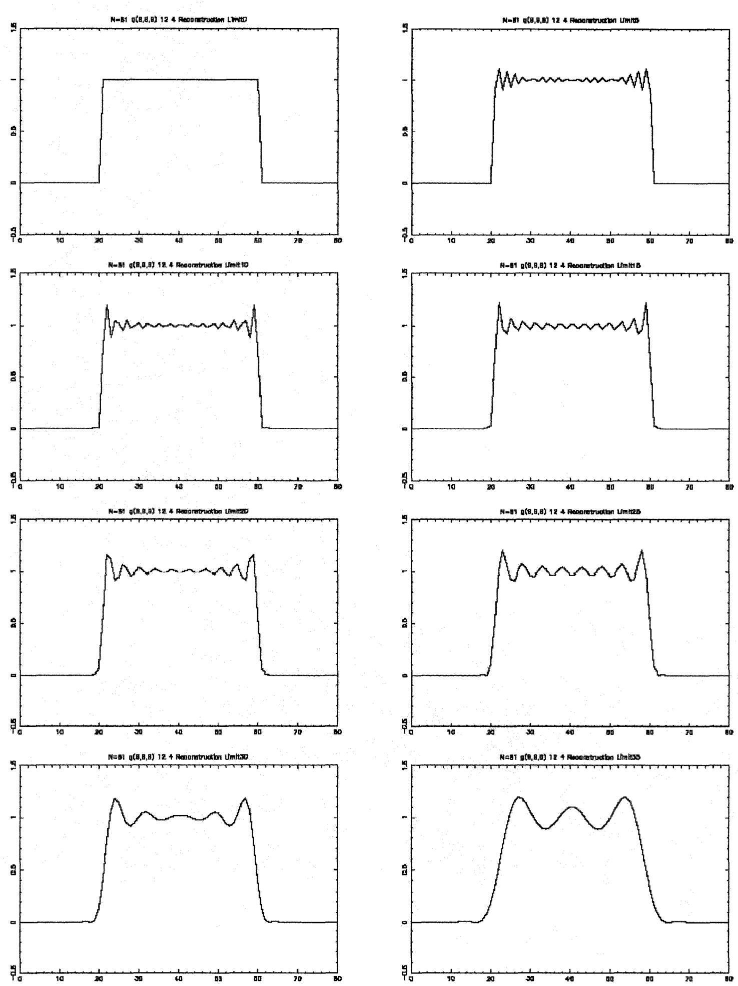

Figure 6: A top hat of was propagated to the Fourier plane where decreasingly small apertures were simulated. Propagation of the limited field was then performed to the next Fourier plane. Each plot is titled with an index Limit# where # gives the number of samples set to zero, to simulate the aperture, from each side of the signal. Qualitatively the throughput of the system decreases from top to bottom and left to right.

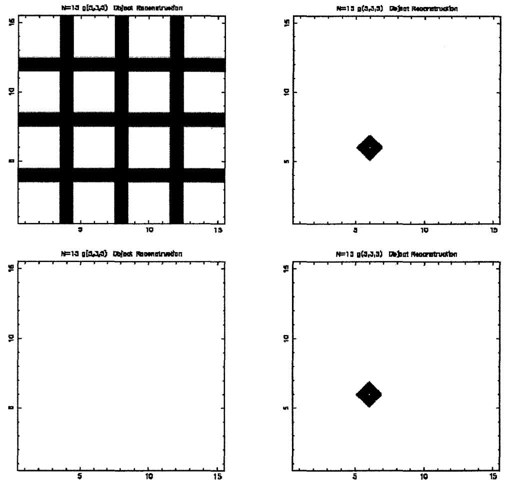

Figure 7: The top left image shows the basis function reconstruction of the intensity at the input plane and the top right image shows a contour map of the coherence function of the array at the same plane centred on a single pixel. Each bolometer element is made up of 9 mutually incoherent pixels and each element is also mutually incoherent. The bottom left plot shows the intensity at the Fourier plane and the bottom right the coherence function at the same plane centred centred on a single pixel.

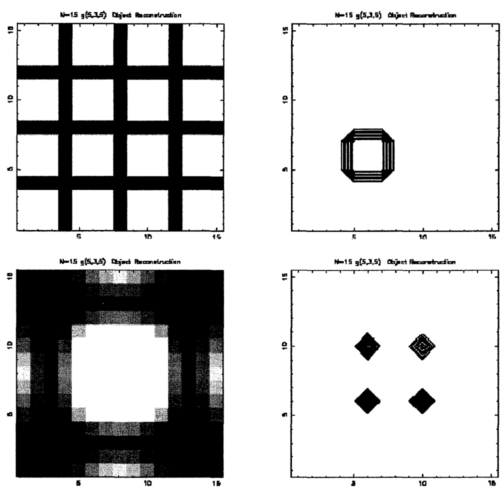

Figure 8: The top left image shows the basis function reconstruction of the intensity at the input plane and the top right image shows a contour map of the coherence function of the array at the same plane centred on a single pixel. Each horn element is made up of 9 mutually coherent pixels (i.e. each element self coherent) but each element is mutually incoherent. The bottom left plot shows the intensity at the Fourier plane and the bottom right the coherence function at the same plane centred centred on a single pixel.

# References


[1] S. Withington and G. Yassin. Power coupled between partially coherent vector fields in different states of coherence. *Optical Society of America Journal*, 18:3061–3071, December 2001.

[2] S. Withington and J.A. Murphy. Modal Analysis of Partially-Coherent Submillimetre-Wave Quasioptical Systems. In *Eighth International Symposium on Space Terahertz Technology*, pages 446–+, 1997.

[3] S. Withington, J. A. Murphy, and G. Yassin. Gaussian-beam-mode analysis of multibeam quasi-optical systems. *IEEE Trans. Antennas and Propagation*, AP:49, 2000.

[4] S. Withington, M.P. Hobson, and R.H. Berry. *To be published*, May 2003.

[5] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36:961–1005, September 1990.

[6] H.P. William, A.T. Saul, T.V. William, and P.F. Brian. *Numerical Recipes in C*. Cambridge University Press, 2 edition, 1992.

[7] S. Qiu and H.G. Feichtinger. Structure of the Gabor matrix and efficient numerical algorithms for discrete Gabor expansions. In *Proc. SPIE Vol. 2308, p. 1146-1157, Visual Communications and Image Processing '94, Aggelos K. Katsaggelos; Ed.*, volume 2308, pages 1146–1157, September 1994.

[8] H.G. Feichtinger, N. Kaiblinger, and P. Prinz. POCS approach to Gabor analysis. In *Proc. SPIE Vol. 3346, p. 18-29, Sixth International Workshop on Digital Image Processing and Computer Graphics: Applications in Humanities and Natural Sciences, Emanuel Wenger; Leonid I. Dimitrov; Eds.*, volume 3346, pages 18–29, January 1998.

[9] D. Gabor. Theory of communication. *J. IEE (London)*, 93:429–457, 1946.

[10] M.J. Bastiaans. Gabor's signal expansion and degrees of freedom of a signal. *Proc. IEEE*, 68:538–539, April 1980.