CHAPTER 4 : PHASE RETRIEVAL WITH THE FAST FOURIER TRANSFORM

4.1 INTRODUCTION AND OVERVIEW

In the previous chapter, we developed a generalised notion of phase which is well-defined for a broad class of partially-coherent free-space radiations such as visible light, X-rays, electrons and neutrons. In the present chapter, we invert a paraxial form of the associated partially-coherent transport-of-intensity equation in order to develop a rapid and stable deterministic algorithm for unique non-interferometric phase recovery.

The problem of deterministic phase retrieval for the case of non-uniform intensity paraxial coherent illumination has been previously tackled using a propagation-based approach ; section 4.2 is a critical evaluation of this earlier method (published by T.E.Gureyev and K.A.Nugent in 1996). We point out that the computing power needed by this previous algorithm makes its use prohibitive unless one has a supercomputer. Furthermore, this approach contains no treatment of partially-coherent illumination[162].

Section 4.3 takes, as a starting point, the partially-coherent transport-of-intensity equation developed in the previous chapter. Assuming the absence of vortices and adopting the paraxial approximation, we develop a symbolic solution for this equation which permits quantitative propagation-based phase imaging of paraxial vortex-free partially-coherent radiation.

Section 4.4 employs this symbolic solution to develop a rapid, determin-istic and stable propagation-based phase retrieval algorithm using the Fast Fourier Transform. This algorithm is tested on simulated data for the case of both coherent and partially-coherent illumination. We also present an analysis of its stability properties with respect to noise.

4.2 OUTLINE AND CRITIQUE OF AN EARLIER METHOD

Prior to undertaking the research outlined in this chapter, a method was developed by T.E.Gureyev and K.A.Nugent for the propagation-based phase retrieval of coherent paraxial scalar electromagnetic radiation[163]. This work

[162] Compare, however, T.E.Gureyev and S.W.Wilkins, *On X-ray phase retrieval from polychromatic images*, Opt.Comm. **147** 229-232 (1998) & *Erratum*, Opt.Comm. **154** 391 (1998) ; T.E.Gureyev, *Transport of intensity equation for beams in an arbitrary state of temporal and spatial coherence* (1998), submitted.

[163] T.E.Gureyev and K.A.Nugent, *Phase retrieval with the transport-of-intensity equation. II. Orthogonal series solution for nonuniform illumination*, J.Opt.Soc.Am.A **13** 1670-1682 (1996) ; T.E. Gureyev and K.A.Nugent, *Rapid quantitative phase imaging using the transport of intensity equation*,

developed a unique and stable deterministic solution to the problem, based upon an expansion of the phase into a weighted sum of orthogonal polynomials. We summarise this earlier method and then point out that the extreme demands it makes on computer hardware are prohibitive for all but the coarsest arrays, unless one makes use of a powerful supercomputer.

(a) Outline of Gureyev-Nugent method

Suppose that we have coherent paraxial scalar electromagnetic radiation incident upon a rectangular aperture, and suppose that we measure the intensity and intensity derivative of the radiation over the aperture. We have already mentioned that, if there are no intensity zeroes over the aperture, then the coherent transport-of-intensity equation yields a unique solution for the phase over the aperture plane[164]. The method of Gureyev and Nugent is a means for obtaining this solution. We briefly summarise this method, a full derivation of which is given in appendix A.

The scaled intensity derivative $F \equiv k\partial_z I$ on the right-hand-side of the coherent paraxial transport-of-intensity equation[165] may be expressed as a weighted sum of Fourier harmonics $W_{mn}(x,y) = \exp(2\pi imx/a)\exp(2\pi iny/b)$, which are orthonormal and complete over the rectangular domain $0 \le x \le a, 0 \le y \le b$. Suppose the Fourier coefficients F_{mn} of the scaled intensity derivative's Fourier expansion $F(x,y) = \sum_{m,n} F_{mn}W_{mn}(x,y)$ are lined up in a column vector. Gureyev and Nugent showed that this column vector of Fourier coefficients F_{mn} can be multiplied by the intensity-dependent matrix $[M_{mn}^{m'n'}]^{-1}$:

$$(1)\quad \phi_{m'n'} = ab\sum_{m,n}\left[M_{mn}^{m'n'}\right]^{-1}F_{mn}$$

to obtain a column vector containing the coefficients ϕ_{mn} in the Fourier expansion $\phi(x,y) = \sum_{m,n}\phi_{mn}W_{mn}(x,y)$ of the phase over the region of interest. The matrix $[M_{mn}^{m'n'}]^{-1}$ is obtained by inverting the matrix $M_{mn}^{m'n'}$:

$$(2)\quad M_{mn}^{m'n'} = 4\pi^2\left(\frac{m'mb}{a} + \frac{n'na}{b}\right)I_{n-n'}^{m-m'},$$

where I_n^m are Fourier coefficients in the expansion $I(x,y) = \sum_{m,n} I_{mn}W_{mn}(x,y)$ of the intensity in the plane $z = 0$.

(b) Critique of Gureyev-Nugent method

The major disadvantage of the method outlined above is that the inversion of the intensity-dependent matrix $M_{mn}^{m'n'}$ is extremely demanding on

---

Opt.Comm. **133** 339-346 (1997). C/f the similar approach in D.Van Dyck and W.Coene, *A new procedure for wave function restoration in high resolution electron microscopy*, Optik **77** 125-128 (1987).
[164] See article 3.3.
[165] See equation (45), article 2.4.2.

computer memory. To retrieve Fourier coefficients ϕ_{mn} up to order M,N one needs to solve a linear system of equations involving a complex matrix of linear dimension 4MN + 2M + 2N elements[166]. The number of floating-point operations required to do this is approximately equal to one third of the cube of its linear dimension (for Gauss-Jordan elimination, which does not require one to calculate the inverse matrix in order to solve the linear system)[167]. Since the Fast Fourier Transform (FFT) of an $M \times N$ pixel array requires $MN\log_2(MN)$ operations, the number of floating-point operations to retrieve the phase to order M,N is[168] $\frac{1}{3}(4MN + 2M + 2N)^3 + 3MN\log_2(MN)$. The memory required to store the array is $2b(4MN + 2M + 2N)^2$ bits, where b is the number of bits of information required to encode each real number. Consider, for example, a 256×256 pixel 16-bit image array. Here, $M_{mn}^{m'n'}$ is a 263168×263168 pixel complex array which would require 2200 gigabytes of memory to store ; the linear system of equations in the Gureyev-Nugent method requires 6×10^{15} floating-point operations to solve. Assuming sufficient memory which can be instantaneously accessed, a CRAY T3E-900 Series supercomputer with a maximum peak performance of 1.8×10^{12} floating-point operations per second would take approximately one hour to recover the phase.

It is clear that the procedure just outlined is very demanding on both computer memory and speed. The remainder of the chapter will be devoted to the development and testing of an alternative algorithm, valid for both coherent and partially-coherent radiation, which substantially reduces demands on computer hardware and makes the phase retrieval of images of substantial size quite accessible to analysis on a modest modern personal computer.

4.3 SYMBOLIC SOLUTION FOR THE PARTIALLY-COHERENT TRANSPORT-OF-INTENSITY EQUATION

We develop a symbolic solution for the partially-coherent transport-of-intensity equation which, in the paraxial approximation, gives the scalar and vector phases associated with our partially-coherent wavefield.

We know from equation (23) of Chapter 3 that the unambiguous part $\vec{S}'_{ave}(\vec{r})$ of the time-average energy flow vector may be written as the gradient of a scalar function $\psi(\vec{r})$; since $\vec{S}'_{ave}(\vec{r})$ is divergence free, we conclude that :

(3) $\nabla^2\psi(\vec{r}) = 0$.

---

[166] T.E.Gureyev and K.A.Nugent (1997).

[167] Information regarding the number of floating point operations required for solving a linear system (without needing to calculate the inverse matrix) and performing FFTs is taken from W.H.Press, S.A.Teukolsky, W.T.Vetterling and B.P.Flannery, *Numerical Recipes in FORTRAN : the art of scientific computing (second edition)*, Cambridge University Press (1992).

[168] The operations used to construct the matrix in (2) are ignored.

Adopt the paraxial approximation and split (3) into contributions from transverse (subscript "$\perp$") and longitudinal components (subscript "z") :

$$(4)\quad \nabla_\perp^2\psi + \partial_z\left(\partial_z\psi\right) = 0\ ,$$

where the dominant direction of energy flow is in the z direction. Next, recall from the discussions centred around equations (16) and (17) of chapter 3, that in both the paraxial and geometric-optics approximations the modulus of the energy-flow vector is proportional to the time-averaged intensity. Since $\vec{S}'(\vec{r}) = \nabla\psi(\vec{r})$ and the radiation is paraxial, we infer that the z-component of the energy-flow vector is proportional to the intensity :

$$(5)\quad \frac{\overline{\omega}k}{4\pi}I\left(\vec{r}_\perp,0\right) \approx \partial_z\psi\left(\vec{r}_\perp,0\right).$$

If we substitute (5) into (4), we may then write down the following formal expression for $\psi(\vec{r})$ in the plane $z = 0$:

$$(6)\quad \psi \approx -\frac{\overline{\omega}k}{4\pi}\nabla_\perp^{-2}\partial_z I\ .$$

We have now recovered ψ in a two-dimensional plane, from which we now obtain an expression for the unambiguous part $\vec{S}' = \nabla\psi(\vec{r})$ of the energy flow vector in $z = 0$. To do this, take the three-dimensional gradient of equation (6):

$$(7)\quad \vec{S}' = \nabla\psi \approx -\frac{\overline{\omega}k}{4\pi}\left(\nabla_\perp + \hat{z}\partial_z\right)\nabla_\perp^{-2}\partial_z I = -\frac{\overline{\omega}k}{4\pi}\left(\nabla_\perp\nabla_\perp^{-2}\partial_z I + \hat{z}\nabla_\perp^{-2}\partial_z^2 I\right)\ .$$

Therefore we can reconstruct the vector field $\vec{S}'$ in the plane z = constant, given knowledge of I, $\partial_z I$ and $\partial_z^2 I$ in that plane.

Rather than reconstructing the vector field associated with the unambiguous part of the energy flow, we are interested in recovering the scalar phase of the wave. To this end, consider the two-dimensional gradient of (6) :

$$(8)\quad \nabla_\perp\psi \approx -\frac{\overline{\omega}k}{4\pi}\nabla_\perp\nabla_\perp^{-2}\partial_z I\ .$$

Next, note from equations (23) to (25) of Chapter 3 that :

$$(9)\quad \nabla\psi\left(\vec{r}\right) = I\left(\vec{r}\right)\left(\nabla\phi_S\left(\vec{r}\right) + \nabla\times\vec{\phi}_V\left(\vec{r}\right)\right),$$

Combining the previous two equations, we have :

$$(10)\quad \nabla_\perp\phi_S\left(\vec{r}\right) + \left[\nabla\times\vec{\phi}_V\left(\vec{r}\right)\right]_\perp \approx -\frac{\overline{\omega}k}{4\pi}\frac{\nabla_\perp\nabla_\perp^{-2}\partial_z I}{I\left(\vec{r}\right)}\ .$$

Next, we take the two-dimensional divergence of (10) :

$$(11)\quad \nabla_\perp^2 \phi_S(\vec{r}) + \nabla_\perp \bullet \left[\nabla \times \vec{\phi}_V(\vec{r})\right]_\perp \approx -\frac{\overline{\omega} k}{4\pi} \nabla_\perp \bullet \left(I^{-1} \nabla_\perp \nabla_\perp^{-2} \partial_z I\right).$$

Let us discard the rotational component (vector phase) in the second term of (11), an excellent approximation when each monochromatic component in the spectral decomposition of the paraxial field is continuous (ie. contains no topological defects). Adopting this approximation, we may then write down the following symbolic solution :

$$(12)\quad \phi_S \approx -\overline{k} \nabla_\perp^{-2}\left[\nabla_\perp \bullet \left(I^{-1} \nabla_\perp \nabla_\perp^{-2} \partial_z I\right)\right].$$

This is the desired symbolic expression for obtaining the scalar phase of a partially-coherent paraxial field with non-uniform illumination.

We are ready to proceed with numerical and experimental investigations based on an implementation of (12) using which utilises the Fast Fourier Transform.

4.4 PHASE RETRIEVAL WITH THE FAST FOURIER TRANSFORM

4.4.1 Phase retrieval algorithm[169]

With a view to the computational implementation of (12), let us first write it in the more convenient form :

$$(13)\quad \phi_S = \phi_S^{(x)} + \phi_S^{(y)}, \quad \begin{cases} \phi_S^{(x)} = -k\left(\nabla_\perp^{-2} \partial_x I^{-1} \partial_x \nabla_\perp^{-2}\right) \partial_z I \\ \phi_S^{(y)} = -k\left(\nabla_\perp^{-2} \partial_y I^{-1} \partial_y \nabla_\perp^{-2}\right) \partial_z I \end{cases}.$$

This may be coded using Fast Fourier Transforms, because the derivative and inverse Laplacian operators become multiplicative operators when acting on the Fourier representation of a function[170] :

$$(14)\quad \nabla_\perp^{-2} = -F^{-1} k_r^{-2} F \quad , \quad \partial_x = iF^{-1} k_x F \quad , \quad \partial_y = iF^{-1} k_y F \quad .$$

Here, F denotes Fourier transformation, F^{-1} denotes inverse Fourier transformation, (k_x,k_y) are the Fourier variables conjugate to (x,y), and $k_r^2 = k_x^2 + k_y^2$. (Note that the Fourier representation of the inverse Laplacian in equation (14) is an example of the pseudo-differential operators discussed in the appendix to chapter 2.) Bearing (14) in mind, the form of (13) which was implemented computationally is given by :

---

[169] This algorithm is covered by an Australian Provisional Patent, *Determination of the phase of a radiation wavefield*, taken out in the names of David Paganin, Anton Barty and Keith Nugent (1998).
[170] We use the Fourier-transform convention given in equation (22) of chapter 2.

$$(15)\quad \phi = \phi^{(x)} + \phi^{(y)}, \quad \begin{cases} \phi^{(x)} = F^{-1} k_r^{-2} k_x F \, I^{-1} F^{-1} k_x k_r^{-2} F \left[k \partial_z I \right] \\ \phi^{(y)} = F^{-1} k_r^{-2} k_y F \, I^{-1} F^{-1} k_y k_r^{-2} F \left[k \partial_z I \right] \end{cases}.$$

It is understood that in the implementation of (15) :

- One will only divide by the intensity if that intensity is greater than a certain threshold value (eg. 0.1% of the maximum value) ;
- Division by k_r^2 does not take place at the point $k_r = 0$ of Fourier space ; instead we multiply by zero at this point. This is related to taking the Cauchy principal value of the integral operator ∇^{-2} (c/f equation (A3) in the appendix to chapter 2).

The speed of the Fast Fourier transform, which may be further enhanced with dedicated hardware, makes (15) an extremely rapid phase retrieval algorithm.

4.4.2 Computer simulations using coherent illumination

Figure 1 gives an example of the action of this algorithm on simulated noise-free coherent data. Diffraction patterns are calculated using the angular-spectrum formalism outlined in article 2.3.3(a). Dimensions of all images are 1.00 cm square = 256×256 pixels. The wavelength of the light was taken to be 632.8 nm (HeNe laser), with defocus distance ± 2 mm. The intensity in the plane $z = 0$, which varies from 0 to 1 in arbitrary units, is given in (a)[171]. Within the area of nonzero illumination, the minimum intensity was 30% of the maximum intensity. (The black border around the edge of the intensity image corresponds to zero intensity.) The input phase, which varies from 0 to π, is shown in (b). The negatively and positively defocused images are given in (c) and (d) respectively, and have respective maximum intensities of 1.60 and 1.75 arbitrary units ; the propagation-induced phase contrast is clearly visible in each of these images. Such propagation-induced phase contrast, well known to microscopists who use defocus to qualitatively visualise phase information[172], is a simple consequence of the local redirection of energy flow as the light passes through the phase object. We see that this propagation-induced phase

[171] Figures 1(a) and 1(b) are courtesy of Public Domain Images, http://www.PDImages.com/.
[172] See, for example, the opening sentence of F.Zernike, *Phase contrast, a new method for the microscopic observation of transparent objects*, Physica **9** 686-693 (1942). This reads, "Every microscopist knows that transparent objects show light or dark contours under the microscope in different ways varying with defocus...". A mathematical formulation of this idea was given in H.Bremmer, *On the asymptotic evaluation of diffraction integrals with a special view to the theory of defocusing and optical contrast*, Physica **18** 469-485 (1952). The final paragraph of Bremmer's paper cites Zernike and gives what is now called the "transport-of-intensity equation" for the case of uniform illumination, noting that "This formula indicates how special features of an object at z = 0 which is invisible there ... become visible for $z \neq 0$, an explanation of which has been given by Zernike."