Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 24 May 2004
Accepted 5 April 2005

Diffraction with wavefront curvature: a path to unique phase recovery

K. A. Nugent,[a]* A. G. Peele,[b] H. M. Quiney[a] and H. N. Chapman[c]

[a]School of Physics, The University of Melbourne, Victoria 3010, Australia, [b]Department of Physics, La Trobe University, Bundoora, Victoria 3086, Australia, and [c]Lawrence Livermore National Laboratory, PO Box 808, Livermore, CA 94550, USA. Correspondence e-mail: keithan@unimelb.edu.au

Modern X-ray optics is able to produce very tightly focused beams. The size of these focused spots is comparable to the scale of large molecules and therefore to the lattice spacing of crystals of these molecules. In this case, the phase of the illuminating beam may vary on the scale of the lattice and conventional diffraction theory needs to be modified. In this paper, coherent diffraction by non-planar beams is considered and it is shown that it is possible to uniquely recover the phase of the diffraction pattern.

1. Introduction

X-ray free-electron lasers are currently the subject of considerable research and development. The applications of these devices are numerous but one of the primary aims is the development of techniques able to probe biomolecular structure using very small crystals, clusters of molecules or even single molecules (Neutze *et al.*, 2000). The molecular structure will be determined from the diffraction patterns produced when molecules pass through the beam.

It has been shown that the diffraction pattern produced by a finite object specifies the object almost uniquely, with there being negligible likelihood of an incorrect structure precisely reproducing the diffraction pattern (Bates, 1982). Current ideas require that the diffracting structure be recovered by iterating between real space and reciprocal space, imposing the measured amplitude in reciprocal space and the known support (spatial extent) of the object in real space (Fienup, 1982) until a consistent solution is found. If the fit to the measurements is sufficiently good, then the uniqueness results quoted above (Bates, 1982) can be invoked to conclude that the correct solution has been found. These methods have been experimentally demonstrated for simple objects (Miao *et al.*, 1999; Robinson *et al.*, 2001; Marchesini *et al.*, 2003) and plausible recoveries have been demonstrated for complex biological objects (Miao *et al.*, 2003). Objective criteria that would establish the reliability of the solution have, however, yet to be found (Elser, 2003) and it is our experience that, if the object has complex interior structure, the object shape will be properly recovered but the interior structure can be unreliable.

In recent years, phase measurement methods have been developed and implemented (Nugent *et al.*, 1996; Cloetens *et al.*, 1996) that use the fact that phase information is carried in the way in which the intensity varies on propagation; these can be termed propagation-based methods. One method that has found wide application is based on a form of the conservation of energy condition known as the transport of intensity equation (Teague, 1983), the solution of which is unique in many circumstances (Gureyev *et al.*, 1995). Other forms of propagation-based phase recovery are based on ideas from electron microscopy using through-focal series of images (Coene *et al.*, 1992). The uniqueness of the phase recovery for the through-focal-series method is an open question but simulations and experiments suggest that the results are generally reliable. Propagation-based methods are applied for real-space phase recovery, such as for optical- (Barty *et al.*, 1998) or electron-microscope images (Bajt *et al.*, 2000), and inherently require that the wavefield contain some curvature.

Prima facie, it would appear that propagation-based methods cannot be applied generally in the context of diffraction experiments in which wavefront curvature is negligible. Modern X-ray optics is progressing very quickly, with the availability of ever increasing amounts of coherent flux and the development of optical elements able to provide excellent spatial resolution. In parallel, modern structural biology is seeking to solve structures with very large unit cells of the order of tens of nanometres, many with a scale comparable to the focal size of a state-of-the-art X-ray optic.

The purpose of the present paper is to consider further whether the convergence of these distance scales allows some new opportunities in diffraction physics. The ability to create and control the phase curvature over small scales offers diffraction physics a new degree of freedom that may be exploited in structural determination (Nugent *et al.*, 2003) and may enable stronger convergence and unique structural determination. In this context, we find that the ideas of propagation-based phase retrieval acquire a natural place in far-field diffraction problems.

2. Overview of methods to retrieve a real-space object from diffraction data

The non-crystallographic phase problem is concerned with the recovery of structural information diffracted by a non-periodic object that is known to be limited in spatial extent. Theoretical descriptions implicitly assume that the incident wave is planar, and that the diffracted intensity is measured in the far-field. The problem of structural recovery is therefore posed in terms of finding the phase of a Fourier transform from its power spectrum.

A number of iterative projection methods based on refinements of the Gerchberg–Saxton algorithm (Gerchberg & Saxton, 1972) have been devised. Discrete Fourier transformations are performed between real-space representations of the object, $f(\mathbf{r})$, and its inverse-space representation, $F(\mathbf{k})$, subject to constraints on the support of $f(\mathbf{r})$ and on the magnitude of $F(\mathbf{k})$, which is fixed at $F(\mathbf{k}) = [G(\mathbf{k})]^{1/2}$ using the diffraction intensity data, $G(\mathbf{k})$. Subsidiary constraints, such as the positivity of $f(\mathbf{r})$, or its external shape, may be imposed provided that one samples $F(\mathbf{k})$ at twice the Nyquist frequency of the object, as discussed by Bates (1982). This is termed oversampling and represents a formal bound on the minimal sampling rate. In practice, a smaller effective oversampling rate combined with some additional *a priori* information is often sufficient to secure a reliable reconstruction of an object by iterative inversion of diffraction data (Miao *et al.*, 1998). These iterative projection algorithms differ mainly in procedural details relating to the determination of $f(\mathbf{r})$ and the phase of $F(\mathbf{k})$, and the imposition of subsidiary constraints. Methods now in common usage include the 'shrink-wrapping' of object shape information (Marchesini *et al.*, 2003), 'charge-flipping' object positivity constraints in crystallographic reconstructions (Oszlányi & Sütő, 2004) and superposition electron-density models (Elser, 2003). Many of these existing methods may be defined within a general projection formulation (Elser, 2003) in which the well known Gerchberg–Saxton and hybrid input–output algorithms arise as special cases.

In the determination of $f(\mathbf{r})$ from $G(\mathbf{k})$, the main practical issues are the usually slow rate of convergence of the iterative projection algorithms, the possible stagnation of the iterative process and the robustness of the algorithms with respect to noise. A substantial literature exists that examines the relative performance of these algorithms under a range of circumstances (Fienup, 1982, 1993; Marchesini, 2004), indicating that the form of the internal features of $f(\mathbf{r})$ has a significant influence on the design and success of the method. A guiding principle is that maximal use should be made of known information about $f(\mathbf{r})$ in an attempt to reduce the number of plausible structures that are consistent with the diffraction data. This information may include some estimate of the size of the object, whether the object is simply connected, real or complex, or constrained in its amplitude or phase structure in some other way.

The augmentation of the known information about an object to enhance its image is a very well established strategy in colour photography. Maxwell demonstrated in the very earliest successful experiments in this field in 1861 that a colour image may be obtained by superimposing three monochrome images of an object, each taken *separately* through an appropriate filter. The reconstruction algorithm in this case simply superimposes these images by projection through matching coloured filters. In an earlier article (Nugent *et al.*, 2003), an algorithm was described that contains an echo of Maxwell's colour photography; three *separate* sets of diffraction data are obtained by varying the magnitude and direction of the cylindrical curvature of the radiation incident on the sample. A detailed specification of $F(\mathbf{k})$ is obtained by processing these data using iterative reconstruction techniques to find the unique $f(\mathbf{r})$ consistent with all three sets.

Finally we note that all far-field single-intensity plane algorithms lose some information about $f(\mathbf{r})$. If we obtain a reconstruction of $f(\mathbf{r})$ that correctly generates $G(\mathbf{k})$ using any phase-retrieval algorithm, then $f(\mathbf{r}+\mathbf{r}_0)$, $f(\mathbf{r})\exp(i\vartheta_0)$ and $f^*(-\mathbf{r})$ satisfy the same equations for a given set $G(\mathbf{k})$, where $\mathbf{r}_0$ is an arbitrary uniform displacement and ϑ_0 is a constant phase shift; the desired solutions are invariant under uniform translation and spatial inversion, and contain an indeterminant absolute phase. These correspond to the 'trivial cases' discussed by Bates (1982).

A further potential uncertainty exists in the case of objects that belong to the class of homometric structures. Two isolated object distributions, $f(\mathbf{r})$ and $g(\mathbf{r})$, may be convolved together to form a third object, $h(\mathbf{r})$, where

$$h(\mathbf{r}) = f(\mathbf{r}) \otimes g(\mathbf{r}). \tag{1}$$

If we further require that $g(\mathbf{r}) \neq g^*(-\mathbf{r})$, then we can form a different function

$$h'(\mathbf{r}) = f(\mathbf{r}) \otimes g^*(-\mathbf{r}). \tag{2}$$

The diffraction patterns of the objects $h(\mathbf{r})$ and $h'(\mathbf{r})$ are clearly identical and cannot be distinguished using oversampling methods. In this paper, we demonstrate that the breaking of symmetry that accompanies the use of astigmatic diffraction techniques enables reconstruction algorithms to distinguish between all of the above-mentioned ambiguities.

3. Focused beams and X-ray optics

Historically, X-ray diffraction theory has been based on an assumption that the incident field is planar, where we here define planar as meaning that the phase distortion over a unit cell is negligible. However, modern crystallography is concerned with structures with unit cells that are increasing in size and, simultaneously, X-ray sources and optics have improved to the point where a tightly focused X-ray beam has a dimension comparable to the size of a large unit cell of a biomolecular crystal. It is therefore timely to re-visit diffraction with this observation in mind. We here consider the phase curvature in the vicinity of a focal spot.

Consider a sample illuminated by, for example, a zone plate. In the case of a perfect zone plate illuminated with perfectly coherent light, the sample will be illuminated with a diffrac-

tion limited spot. The phase distribution at the focus is flat but rapidly acquires curvature away from the focus, maximizing at the so-called Rayleigh distance (Saleh & Teich, 1991), and then reducing again to zero curvature in the far-field. To quantify this, let us consider Gaussian optics. Standard results tell us that a converging beam with $1/e^2$ waist w_0 has a minimum radius of curvature given by $R = 2\pi w_0^2/\lambda$. A maximum phase deviation from planar of $\Delta\Phi = 2\pi p$ is therefore obtained with a Gaussian beam waist of $w_0 = d_{\rm obj}/4(\pi p)^{1/2}$. In the case of $\Delta\Phi \simeq 1/8$ (*i.e.* $p = 1/16\pi$), the required spot size is $w_0 \simeq d_{\rm obj}$. That is, when the focal spot is comparable in size to the size of the object (or the unit cell in the crystal), the wavefront curvature over the object will be significant when it is placed at the Rayleigh distance.

Modern zone-plate technology combined with third-generation synchrotron sources can deliver focal-spot sizes of around 100 nm for harder X-rays (di Fabrizio *et al.*, 1999) and 30 nm or better (Kipp *et al.*, 2001) for softer X-rays, which suggests that it should be possible to introduce significant phase curvature for crystal structures with a unit cell with a size greater than 30 nm using already existing technology. This brings the method into the range of moderate-sized protein crystals. It is to be anticipated that 10 nm will be achieved in the foreseeable future, although it has been claimed that the technology may go no further for rather fundamental reasons (Bergemann *et al.*, 2003).

4. Unique phase recovery using coherent astigmatic diffraction

4.1. General theory

In an earlier article (Nugent *et al.*, 2003), the effect of a small phase change, $\alpha(\mathbf{r})$, was considered for an otherwise conventional diffraction experiment, where $\mathbf{r}$ is a specified position in a two-dimensional plane perpendicular to the direction of propagation of the incident beam. It was further assumed that the sample is sufficiently thin that the effective perturbing potential can be regarded as a function of $\mathbf{r}$ in the sample plane. This requires that the sample width, δz, in the direction of propagation, z, satisfies the condition $\delta z \ll 1/k_{\rm max}$, where $k_{\rm max}$ is the maximum measured wavenumber in the diffraction data.

In the case of a parabolic phase change, $\alpha(\mathbf{r}) = k_0 r^2/2R$, where $k_0 = 2\pi/\lambda$ is the wavenumber of the incident radiation, λ is the corresponding wavelength and R is the characteristic scale of the phase change. The illuminating wave may be written in the form

$$\psi_{\rm inc} \simeq \left[1 + ik_0\frac{r^2}{2R}\right]\exp(-ik_0 z). \tag{3}$$

The leading-order perturbation correction to $\psi_{\rm inc}$ is second-order in r, enabling the transformation of the equation for the effect of the perturbation on the diffracted intensity to a form that depends on the second derivative in the reciprocal-space variable, $\mathbf{k}$.

$$\delta I_f(\mathbf{k}) = \frac{1}{k_0^3 R}\nabla\cdot\{I_f(\mathbf{k})\nabla\Phi_f(\mathbf{k})\}, \tag{4}$$

where $\delta I_f(\mathbf{k}) = I_f(\mathbf{k}) - I'(\mathbf{k})$, $I_f(\mathbf{k})$ represents the diffracted intensity data in the absence of the perturbation, $I'_f(\mathbf{k})$ is the corresponding data in the presence of the perturbation and $\Phi_f(\mathbf{k})$ contains the diffracted phase information. The structure of equation (4) is formally identical to the transport of intensity equation that is used in real-space phase recovery. In this case, however, the fields are described in terms of the reciprocal-space variable, $\mathbf{k}$, and there is no requirement to assume the validity of the paraxial condition.

When posed as a Neuman problem, equation (4) has a unique solution given a knowledge of the boundary condition $\hat{\mathbf{n}}\cdot\nabla\Phi_f = g$, where $\hat{\mathbf{n}}\cdot\nabla\Phi_f$ is the phase gradient normal to the boundary of the simply connected bounded two-dimensional solution domain. This uniqueness condition assumes, however, that $I_f(\mathbf{k}) > 0$, which is not fulfilled for most diffracted fields, and is clearly violated whenever phase discontinuities generate zero-valued points in the intensity. In practice, equation (4) would rarely be associated with a unique solution to the phase-retrieval problem. Nonetheless, as will be seen, it can lead to satisfactory solutions in practice.

4.2. Astigmatic diffraction and uniqueness

The indeterminacies in the phase arise owing to symmetries within the structure of its discontinuities, which one expects to be broken by the external introduction of asymmetry in the optical system. This may be achieved by introducing a cylindrical rather than a spherical wave, corresponding to a phase perturbation $\alpha(\mathbf{r}) = k_0 q^2/2R$, where $q = x$ or $q = y$ denotes the direction of the introduced cylindrical astigmatism.

An argument analogous to that for a spherical perturbation leads to the expression

$$\delta I_f^q(\mathbf{k}) = \frac{1}{k_0^3 R}\frac{\partial}{\partial k_q}\left\{I_f(\mathbf{k})\frac{\partial}{\partial k_q}\Phi_f(\mathbf{k})\right\}, \tag{5}$$

where the superscript label q in δI_f^q indicates that $\delta I_f(\mathbf{k})$ has been obtained using curvature in the q direction. In an earlier paper on this subject (Nugent *et al.*, 2003), the role of the boundary conditions was considered in the context of an object of known finite extent. In this article, we are also concerned with phase recovery for objects of infinite extent illuminated by a focused (or, more generally, finite energy) beam and so we will explore this issue a little further.

In the present case, the Poynting vector in the far field, $S(\mathbf{k}) = (1/k_0^2)I_f(\mathbf{k})\nabla\Phi(\mathbf{k})$, is considered to be a two-dimensional function of $\mathbf{k}$. In the presence of cylindrical phase curvature, $S_q(\mathbf{k})$, the q component of $S(\mathbf{k})$, is defined with respect to the available information by the differential equation

$$\left.\frac{\partial S_q(\mathbf{k})}{\partial k_q}\right|_{k_{\bar{q}}} = \left.k_0 R\delta I_f^q(\mathbf{k})\right|_{k_{\bar{q}}}, \tag{6}$$

where $\bar{q}$ denotes the direction perpendicular to q. For illumination by a localized beam, $I_f(\mathbf{k}) \rightarrow 0$ and $S_q(\mathbf{k})|_{s_{\bar{q}}} \rightarrow 0$ in the

limit $k_q \rightarrow \pm\infty$, which fixes the boundary condition on the solution of the differential equation for $S_q(\mathbf{k})|_{k_{\bar{q}}}$. A unique solution to this equation is thus obtained.

If we consider the function $\delta I_f(\mathbf{k})$, however, we note that, for any fixed value of $k_{\bar{q}}$, the direct integration of equation (6) using, for example, finite difference techniques is likely to encounter a discontinuity in the derivative of $\delta I_f^q(\mathbf{k})$ owing to the presence of phase vortices. The functions $S_q(\mathbf{k})|_{k_{\bar{q}}}$ are consequently only piecewise continuous, and it would be necessary to determine accurately the Fourier space positions of these vortices in any direct numerical attempt to obtain them by solution of equation (6). Such a direct solution is nonetheless possible, at least in principle.

Following Nugent *et al.* (2003), it is easily shown that the Poynting vector uniquely specifies the incident field to within a constant phase term. Suppose that two fields have the same Poynting vector but different phases, so that

$$I(\mathbf{k})\nabla\Phi_1(\mathbf{k}) = I(\mathbf{k})\nabla\Phi_2(\mathbf{k}). \tag{7}$$

For non-zero intensity, one may write

$$I(\mathbf{k})\nabla[\Phi_1(\mathbf{k}) - \Phi_2(\mathbf{k})] = 0, \tag{8}$$

for which the solution is

$$[\Phi_1(\mathbf{k}) - \Phi_2(\mathbf{k})] = \text{constant}. \tag{9}$$

A constant phase difference has no physical meaning and the two fields are consequently identical.

We therefore conclude that a measurement of the far-field diffraction pattern combined with far-field diffraction patterns obtained with orthogonal cylindrically curved waves are sufficient to determine uniquely the phase of the diffraction pattern. We refer to the method as astigmatic diffraction.

5. Periodic structures

The derivation in the previous section assumes that all the derivatives of $I_f(\mathbf{k})$ are defined in order to recover equation (5). This is acceptable when the diffracting object is finite in extent and so its diffracted field is known to be analytic. However, the phase problem for most of crystallography concerns essentially infinite periodic structures and this is now the problem that we consider. We envisage that the principal application of this method will be to small crystals and so, for simplicity, we will retain the Born approximation.

Consider a structure described by

$$\sigma(\mathbf{r}) = \sum_{\mathbf{G}} n_G \exp(i\mathbf{G}\cdot\mathbf{r}) \tag{10}$$

illuminated with an incident wave $\psi_{\text{inc}} = \exp(-i\mathbf{k}_0\cdot\mathbf{r})$. The far-field intensity distribution is given by

$$I(\mathbf{k}) \propto \left|\sum_{\mathbf{G}} n_G \delta(\mathbf{k} - \mathbf{k}_0 + \mathbf{G})\right|^2, \tag{11}$$

which gives the familiar Bragg diffraction condition

$$\mathbf{k} - \mathbf{k}_0 = -\mathbf{G}. \tag{12}$$

This is the basis of conventional diffraction theory. Now consider the crystal illuminated with a field containing the general variation, $\exp[\alpha(\mathbf{r})]$, so that $\psi_{\text{inc}} = \exp(-i\mathbf{k}_0\cdot\mathbf{r})\exp[\alpha(\mathbf{r})]$, where $\alpha(\mathbf{r})$ may contain information about spatial variations in both intensity and phase. In this case, the scattered wave, ψ_{scat}, is given by

$$\psi_{\text{scat}} = \sum_{\mathbf{G}} n_G \exp[i(\mathbf{G} - \mathbf{k}_0)\cdot\mathbf{r} + \alpha(\mathbf{r})] \tag{13}$$

and the far-field intensity is of the phase-modified form

$$I(\mathbf{k}) \simeq \left|\sum_{\mathbf{G}} n_G \int \exp[i(\mathbf{k} - \mathbf{k}_0 + \mathbf{G})\cdot\mathbf{r} + \alpha(\mathbf{r})]\,\mathrm{d}\mathbf{r}\right|^2. \tag{14}$$

In practice, we introduce phase and intensity variation over the sample using curved wavefronts and a sufficiently general representation of $\alpha(\mathbf{r})$ may be written in the form

$$\alpha(\mathbf{r}) = -\lambda_x x^2 - \lambda_y y^2 = \alpha_x(\lambda_x; x) + \alpha_y(\lambda_y; y), \tag{15}$$

where x and y are distances measured perpendicular to the beam axis in the plane of the sample. The parameters λ_x and λ_y are, in general, complex; the real part of λ_q for $q \in \{x, y\}$ describes the intensity variation in a Gaussian approximation if $\Re(\lambda_q) > 0$, while $\Im(\lambda_q)$ defines the phase variation of a spherical wave in a plane perpendicular to the direction of propagation. The phase variation, in particular, typically depends on the radius of curvature of the wavefront, R, and the wavenumber of the incident radiation, k_0, while the sign of $\Im(\lambda_q)$ is determined by whether the target plane is upstream or downstream of the focal plane. Note that in contrast to the development in the previous section there is no longer a requirement in what follows that $\alpha(\mathbf{r})$ be in any sense 'small'. For notational convenience, we define

$$(\mathbf{k} - \mathbf{k}_0 + \mathbf{G})\cdot\mathbf{r} = w_x(G_x, u_x, u_{x,0})x + w_y(G_y, u_y, u_{y,0})y. \tag{16}$$

We may then write equation (14) as

$$I(\mathbf{k}) \simeq \left|\sum_{\mathbf{G}} n_G I_x(w_x, \lambda_x) I_y(w_y, \lambda_y)\right|, \tag{17}$$

where it is to be understood that $w_x = w_x(G_x, u_x, u_{x,0})$, $w_y = w_y(G_y, u_y, u_{y,0})$ and

$$I_q(w_q, \lambda_q) = \int_{-\infty}^{\infty} \exp(iw_q q - \lambda_q q^2)\,\mathrm{d}q \tag{18}$$

for $q = \{x, y\}$. These integrals may be evaluated using the standard result (Abramowitz & Stegun, 1970; Sneddon, 1951)

$$I_q(w_q, \lambda_q) = \left(\frac{\pi}{\lambda_q}\right)^{1/2} \exp\left(\frac{-w_q^2}{4\lambda_q}\right), \tag{19}$$

which is valid for $\Re(\lambda_q) > 0$. In order to examine the analyticity of $I(\mathbf{k})$ in the neighbourhood of $w_q = 0$ and $\lambda_q = 0$, this restriction on λ_q is imposed by writing $\lambda_q = \varepsilon_q - i\gamma_q$, where ε_q is real and positive.

In order to recover conventional diffraction theory and the Bragg condition, we set $\lambda_q = 0$, in which case we obtain

$$I(w_q, 0) = \lim_{\varepsilon\rightarrow 0_+} \left(\frac{\pi}{\varepsilon_q}\right)^{1/2} \exp\left(-\frac{u_q^2}{4\varepsilon_q}\right) = 2\pi\delta(w_q), \tag{20}$$

which regenerates the result embodied in equation (11) for each component, w_q.

In practice, the use of a Gaussian beam profile ensures that ε_q satisfies the conditions required by equation (19), since the Gaussian intensity profile corresponds to a real positive value of ε_q. The diffraction integrals, and hence the diffracted intensity, equation (17), are analytic in the direction of the applied wavefront curvature, excluding at the parametric Bragg limit $w_q \rightarrow 0$. Subject to the assumptions inherent in the development in equation (4), and especially the weak curvature assumption implicit in equation (4), one concludes that the phase recovery yields a unique result even when the diffracting structure is infinite and periodic.

It can further be seen that non-planar incident fields will diffract off reciprocal-lattice vectors that would not otherwise be detected without rotating the crystal and that the curved wavefront will act to spread the diffracting orders out over the detector. If the diffracted spots are sufficiently diffuse, then they will begin to overlap and interference fringes form. The phase of the interference fringes will yield information about the relative phases of the diffracted orders which will itself yield additional structural information. In this limit, the method is related to the ptychographic approach that is used in electron microscopy (Hegerl & Hoppe, 1970; Nellist & Rodenburg, 1998).

6. Solving for the phase

Equation (4) has a formal structure identical to the transport of intensity equation. This equation has been applied extensively in quantitative phase imaging (Bajt *et al.*, 2000). Notwithstanding the formal mathematical difficulties associated with the presence of phase discontinuities, this method is known to provide a robust retrieval algorithm. In this section, we explore the application of equation (4) to far-field diffraction studies.

The astigmatic methods discussed in §4.2 offer the possibility of a formally unique solution. The associated symmetry breaking and increased information available to the algorithm suggest that improved convergence characteristics may be obtained. The two approaches are critically compared in this section.

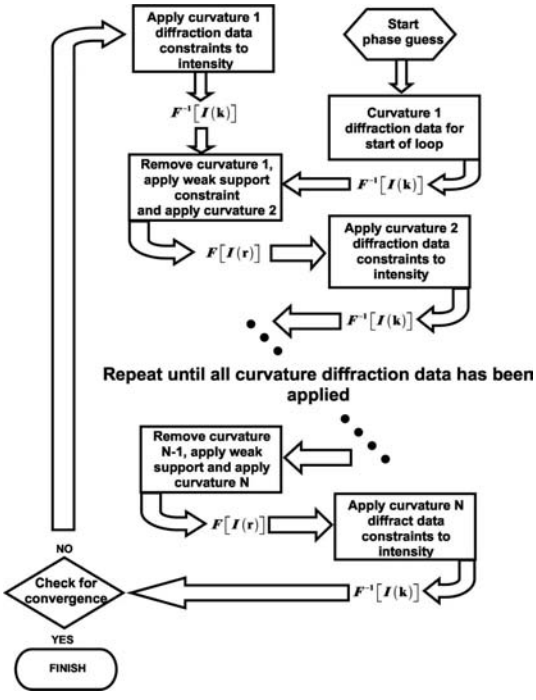

Figure 1
General algorithm for iterative phase reconstruction, which is described in the text in §6.1. When all curvatures are spherical then the algorithm is the iterative form for solution of equation (4). When one or more sets of orthogonal cylindrical curvatures are used then the algorithm is the iterative form for the solution of equation (5).

6.1. Algorithm for phase recovery

Although it would be possible to directly integrate a Poynting vector field to recover the phase distribution, a diffracted field will typically contain many vortices and so this will, in general, not be practical. We therefore developed an iterative scheme to recover the phase. The algorithm is similar in spirit to other iterative methods and can be reduced to the standard Gerchberg–Saxton scheme, as indicated below. To commence the iteration, a phase guess is applied to an initial set of diffraction data. An inverse Fourier transform gives the current object plane solution where a support constraint may be applied. Any curvature used to obtain the diffraction data is then removed and the curvature corresponding to the next set of diffraction data is applied. A Fourier transform then gives the current diffraction plane solution where the actual diffraction intensity is applied. The cycle is repeated through all curvature diffraction data sets and a test for convergence is applied before repeating the iteration. The conventional Gerchberg–Saxton algorithm corresponds to $N = 1$ and vanishing effective curvature. When all curvatures are spherical then the algorithm is the iterative form for solution of equation (4). When one or more sets of orthogonal cylindrical curvatures are used then the algorithm is the iterative form for the solution of equation (5).

In order to assess convergence, we define a quality of fit parameter against the correct solution, determined over the support,

$$R = \frac{\sum_{ij} ||d_{ij}^{\rm recon}| - d_{ij}^{\rm truth}||}{\sum_{ij} ||d_{ij}^{\rm recon}| + d_{ij}^{\rm truth}||}, \qquad (21)$$

where $d_{ij}^{\rm recon}$ is the ijth pixel of the reconstruction and $d_{ij}^{\rm truth}$ is the corresponding correct value. For the model systems adopted here, this definition of R allows us to construct a measure of the quality of the reconstructed target gauged against the known control of the function that has been used to generate the simulated diffraction data. In practical applications, however, one would would need to devise a quality metric that is based solely on comparison of the experimental and reconstructed diffraction data, since the details of $d^{\rm truth}$ will, in general, be unknown.

6.2. Algorithm comparison

In order to assess the performance of some of the different variations possible for the iterative algorithm shown in Fig. 1, we simulated the diffraction pattern of an artificial molecule made of 40 randomly positioned C atoms within a 75 nm square window. The object distribution is shown in Fig. 2. Curved diffraction data were generated with a maximum curvature corresponding to an approximate phase shift of π at the edge of the object. Astigmatic and spherical diffraction data sets were generated. Multiple curved data sets were also generated with the curvature reducing from π for the first data set to $\pi/5$ for the fifth data set in equal steps. Finally, we also investigated the effect on the performance of the algorithm when a constraint was not used at all. This is, in principle, possible where multiple data sets are used as the additional data forces a self-consistent solution.

Plots showing the evolution of the fit quality parameter, R, of equation (21) as a function of iteration number using several variations on the algorithm are shown in Fig. 3. The curves from left to right correspond to the following cases: five sets of x and y cylindrical curvature data with support constraint; five sets of x and y cylindrical curvature data with no constraint; one set of x and y cylindrical curvature data with support constraint; five sets of spherical curvature with support constraint; and five sets of spherical curvature with no constraint. After 400 iterations, there is no appreciable convergence for the Gerchberg–Saxton method and so the corresponding plot is not shown.

In all cases shown, the retrieved object structure was indistinguishable from that shown in Fig. 2. It can be seen for the object examined here that the support constraint offers only a slight improvement in the speed of convergence to the correct solution. In early iterations, the application of the support constraint may even reduce the apparent convergence rate. A direct comparison of initial convergence rate with and without the support constraint must be heavily qualified, however, since the starting guess for the target object is different in the two cases, and the algorithms follow independent and unrelated routes towards convergence. One does observe, however, that the asymptotic convergence is more rapid if the support constraint is applied, compared with the case in which it is absent. The improvement obtained by incorporating multiple diffraction data sets into the algorithm is clear. In practice, many data sets will also serve as a method of reducing experimental error. Finally, we note the clear advantage obtained in using astigmatic diffraction data sets compared with spherical. We conjecture that the use of astigmatic diffraction data sets introduces further information into the iterative system, effectively doubling the number of multiple diffraction data sets available using spherical data alone. In the following subsection, we consider objects that cannot be determined using planar or spherical incident fields.

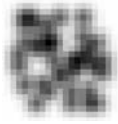

Figure 2
Test object used to investigate the algorithm described in Fig. 1. The recovered structure is visually indistinguishable from the original structure of the test object shown here.

6.3. Vortex phase objects

Astigmatic diffraction was directly simulated for a vortex optical field of the form

$$\psi(\mathbf{r}) = \begin{cases} \exp[-(r/r_0)^2] + im\theta & r \neq 0 \\ 0 & r = 0 \end{cases} \qquad (22)$$

with $\mathbf{r} = r\cos(\theta)\hat{\mathbf{i}} + r\sin(\theta)\hat{\mathbf{j}}$, $m = \pm 1$, $r \geq 0$ and $0 \leq \theta < 2\pi$. The Fourier transform was calculated and it was confirmed that the diffracted intensity was independent of the sign of the topological charge, m. This field therefore has a genuine phase sign ambiguity that cannot, in principle, be resolved using the oversampling method. Small (maximum phase excursion of much less than π) orthogonal cylindrical phase curvatures were introduced and equation (5) was applied to recover the Poynting vector for the field. The result is shown in Fig. 4 and it can be seen that the vortex structure is accurately recovered and the phase gradient circulates around the phase discontinuity at the intensity zero. The reconstructed topological charge was confirmed to change sign with a change in the initial field. The phase recovery for this structure is shown in Fig. 5.

We also performed the reconstruction using the Gerchberg–Saxton method and were sometimes able to obtain solutions that were correct. However, sometimes the solution was reversed in sign and sometimes the iteration failed to converge, thus confirming the non-uniqueness of solutions obtained by non-astigmatic methods.

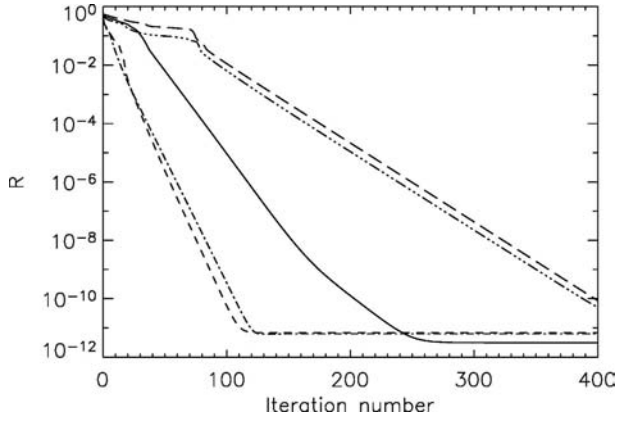

Figure 3
Plots showing the evolution of the quality of fit parameter of equation (21) as a function of iteration number using the algorithm described in Fig. 1. The curves from left to right at $R = 10^{-6}$ correspond to the following cases: five sets of x and y cylindrical curvature data with support constraint; five sets of x and y cylindrical curvature data with no constraint; one set of x and y cylindrical curvature data with support constraint; five sets of spherical curvature with support constraint; and five sets of spherical curvature with no constraint.

6.4. Homometric object

As indicated in §2, a reduction of symmetry in image retrieval algorithms brings with it the ability to distinguish between nearly identical structures or between the elements of homometric classes of structures that are strictly indistinguishable using uniform plane-wave illumination.

Fig. 6 shows the reconstruction of two objects that are homometric. The reconstructions are indistinguishable from the original objects. This example demonstrates that objects of this type are reconstructed correctly to within the numerical accuracy of the computer using astigmatic diffraction data.

6.5. Infinite periodic objects

A previous paper (Nugent *et al.*, 2003) considered diffraction by an isolated object with compact support, and these results have been extended here. It has been shown that unique reconstruction of infinite periodic samples should also be possible. In order to test the reconstruction, homometric periodic structures were calculated (Fig. 7). The structures were assumed to be illuminated with a curved beam to ensure that the diffracted field had well defined spatial derivatives. The data were then entered into the iterative program. It was found that the temporary application of a finite support was helpful to initiate convergence, but once this had begun the support was removed and the algorithm converged on the correct result very smoothly. The finite support was chosen to encompass several unit cells, although the precise selection of this region is not otherwise critical.

6.6. Astigmatic diffraction

The astigmatic diffraction method introduces additional information into the iterative algorithm. The use of astigmatic information creates the ability to generate unique solutions. While these are referred to as 'trivial' cases by Bates, we see from §§6.3 and 6.4 that they can, nonetheless, be intrinsically interesting structures. Moreover, the use of this method leads to the intriguing possibility of solving the structure for periodic objects. The fact that two or more astigmatic diffraction data sets must be used also means that there is more information available to the algorithm than for the conventional Gerchberg–Saxton or spherical curvature methods. Accordingly, not only does astigmatic diffraction offer the security of a unique solution, it gets to that solution faster than other methods.

7. Practical considerations

This paper has argued that the illumination of a sample using a beam containing wavefront curvature can be used to determine uniquely the diffracting structure. A central part of this argument is that the curvature can be sufficiently well known to allow the equations to be solved, that sufficient diffraction can be observed and that the system position is stable enough for multiple data sets to be reliably acquired. These are all important experimental issues. In this section, we briefly consider implementation approaches that might allow multiple curvature diffraction to be applied in practice.

In this work, we have considered the illuminating of samples using a small additional cylindrical curvature added to a curved beam. The appropriate wavefront might be produced by focusing the X-ray beam onto the sample using a conventional zone plate and then introducing an additional weakly

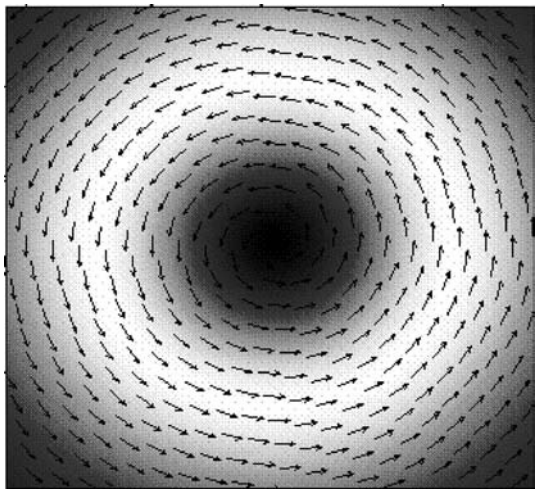

Figure 4
Poynting vector recovered from an optical vortex beam. The handedness of the circulation was always correctly recovered.

Figure 5
Recovery of the vortex phase distribution.

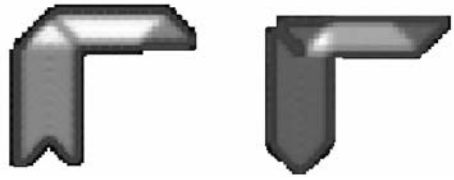

Figure 6
Recovered non-periodic homometric objects. These images show the recovered structures. They are indistinguishable from the original input.

focusing cylindrical lens. An alternative might be to rotate the conventional zone plate about an axis orthogonal to the optic axis so as to introduce astigmatism to the optical system and thereby produce a cylindrically curved component. Another possibility is to use a Kirkpatrick–Baez imaging system and adjust the mirror angle to produce cylindrical curvature. It can be concluded, therefore, that the introduction of appropriate curvature is well within the capabilities of modern X-ray optical technology.

In conventional diffraction physics, the location of the sample within the beam is irrelevant because, by assumption, the incident beam is planar. This allows a great deal of latitude in the alignment of the experimental apparatus. The essential point of this paper is to remove the plane-wave assumption and so this matter must be considered more deeply. Indeed, given the proposals in the previous paragraph, it must be acknowledged that the optical adjustments required to change the wavefront curvature are certain to move the beam with respect to the sample. Suppose that introducing the cylindrical curvature moves the beam laterally a small amount. It is simple to show that such a shift adds an additional linear phase to the diffracted wave. In the far-field, this will correspond to a translation of the intensity pattern. Thus, it is likely that a simple alignment procedure will be required.

If the intensity distribution of the beam is changing significantly relative to the spatial scale of the unit cell then additional complexity will be introduced. In §3, we considered the production of highly curved beams using focusing optics. Such an approach will have two associated complexities. First, the intensity distribution will have a significant spatial variation with respect to the unit cell. Thus, corrections to this effect would have to be explored. This does not seem to represent a difficult correction to make, but should be the subject of further work. Secondly, the region of the diffracting structure being illuminated will be very small and this could result in considerable radiation damage, or a very small diffracted signal.

One might imagine, in the context of an X-ray free electron laser experiment, that the molecular systems (being single molecules, clusters or nanocrystals) might be introduced in different parts of a focusing beam, as illustrated in Fig. 8. The resulting diffraction patterns will then correspond to different incident curvatures. An experiment designed along these lines would represent only a minor increase in complexity over the acquisition of diffraction patterns at one point in the beam.

In order to make use of the phase and intensity structure of a beam in the reconstruction algorithm, one must be able to characterize these features on planes illuminating the target that are perpendicular to the propagation axis of a real beam. The wavefield of a focused beam may, at least in principle, be obtained by phase retrieval of its image in the detector plane using an adaption of the Gerchberg–Saxton algorithm. In order to characterize the illumination of a sample, this wavefield can then be propagated a short distance from the focal plane to the target plane in the Fresnel approximation. Since the function describing the target object and the illuminating wavefield are multiplicative, any small residual error in this step manifests itself as a slowly varying modulation of the reconstruction of the object. This modulation introduces only small errors in the reconstruction while preserving the benefits of utilizing phase curvature as a means to accelerate iterative phase retrieval that are described in the present article.

Finally, the illuminating field will itself be highly divergent and so the diffraction pattern will be partly superimposed on a very bright undiffracted beam. However, the spatial frequencies contained in the region of the undiffracted background would correspond to a spatial scale (*i.e.* a resolution) comparable to that of the beam size. Thus the useful structural information will be scattered out of this beam and so superposition on the divergent undiffracted beam should not, at least in principle, imply a fundamental problem.

Finally, we note that the key requirement for astigmatic diffraction would appear to be the need for waves with complex symmetry properties. It would be interesting to explore the application of well characterized but complex wavefronts with, for example, a large local curvature

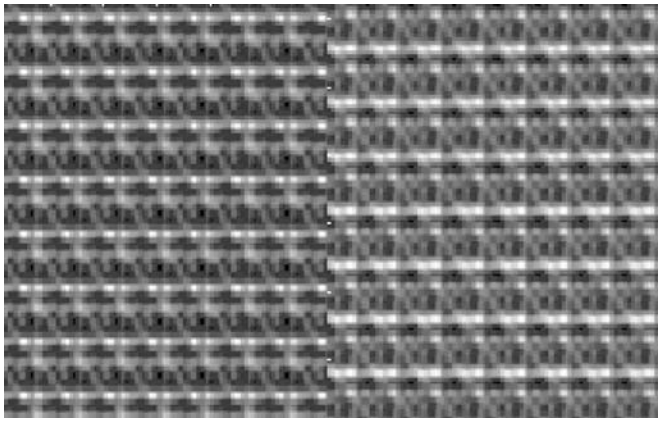

Figure 7
Recovered homometric infinite periodic structures. These are indistinguishable from the input distribution.

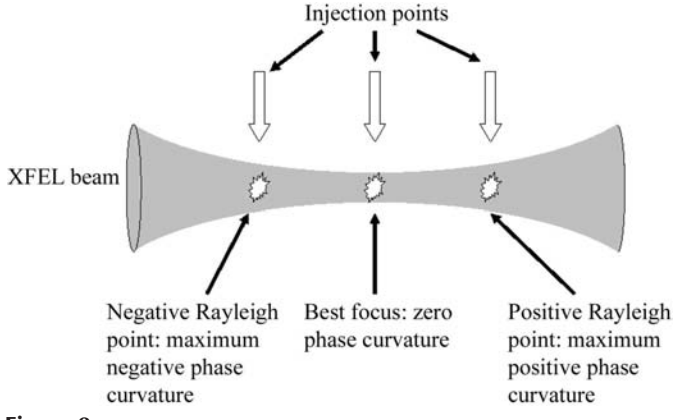

Figure 8
Schematic of a possible free-electron laser experiment. Measurements are obtained of the X-rays diffracted by molecules placed into the beam at different curvature points. A conventional focused beam will produce different spherical curvatures and an astigmatic beam will produce astigmatic diffraction data.

combined with a small overall curvature to the diffraction problem. We speculate that such wavefronts would allow unique phase recovery.

8. Conclusions

We have presented a discussion of the diffraction of curved wavefronts by arbitrary samples. The diffraction model used is rather simple and we acknowledge that it ignores many of the complexities of a full dynamical diffraction theory. Nevertheless, we believe that the method proposed here should be able to yield a practical approach that will allow the unique solution of the phase problem for a broad range of diffracting samples. We have presented simulations that have supported our theoretical conclusions and we have found that rapid convergence on the correct solution is invariably attained. The simulations used are, of course, highly idealized and in §7 we have presented a brief consideration of some of the practical issues that would need to be considered. In summary, we believe that the very highly shaped beams offered by modern X-ray optics offers a hitherto unrecognized flexibility for the acquisition of diffraction data.

The authors acknowledge the support of the Australian Research Council for both grant support and fellowships to AGP and KAN. The work of HNC was performed was performed under the auspices of the US DOE by LLNL under contract no. W-7405-ENG-48. We also thank Jose Varghese of the CSIRO for useful discussions.

References

Abramowitz, M. & Stegun, I. A. (1970). *Handbook of Mathematical Functions.* New York: Dover Publications.
Bajt, S., Barty, A., Nugent, K. A., McCartney, M., Wall, M. & Paganin, D. (2000). *Ultramicroscopy*, **83**, 67–73.
Barty, A., Nugent, K. A., Paganin, D. & Roberts, A. (1998). *Opt. Lett.* **23**, 817–819.
Bates, R. H. T. (1982). *Optik (Stuttgart)*, **61**, 247–262.
Bergemann, C., Keymeulen, H. & van der Veen, J. F. (2003). *Phys. Rev. Lett*. **91**, 204801.
Cloetens, P., Barrett, R., Baruchel, J., Guigay, J. P. & Schlenker, M. (1996). *J. Phys. D*, **29**, 133–146.
Coene, W., Janssen, G., Debeeck, M. O. & van Dyck, D. (1992). *Phys. Rev. Lett*. **69**, 3743–3746.
Elser, V. (2003). *J. Opt. Soc. Am.* **20**, 40–55.
Fabrizio, E. di, Romanato, F., Gentili, M., Cabrini, S., Kaulich, B., Susini, J. & Barrett, R. (1999). *Nature (London)*, **401**, 895–898.
Fienup, J. R. (1982). *Appl. Opt.* **21**, 2758–2769.
Fienup, J. R. (1993). *Appl. Opt.* **32**, 1737–1746.
Gerchberg, R. & Saxton, W. (1972). *Optik (Stuttgart)*, **35**, 237.
Gureyev, T. E., Roberts, A. & Nugent, K. A. (1995). *J. Opt. Soc. Am.* **12**, 1942–1946.
Hegerl, R. & Hoppe, W. (1970). *Ber. Bunsenges. Phys. Chem.* **74**, 1148–1154.
Kipp, L., Skibowski, M., Johnson, R. L., Berndt, R., Adelung, R., Harm, S. & Seemann, R. (2001). *Nature (London)*, **414**, 184–188.
Marchesini, S. (2004). *arXiv:physics*, **1**, 0404091.
Marchesini, S., He, H., Chapman, H. N., Hau-Riege, S. P., Noy, A., Howells, M. R., Weierstall, U. & Spence, J. C. H. (2003). *Phys. Rev. B*, **68**, 140101.
Miao, J. W., Charalambous, P., Kirz, J. & Sayre, D. (1999). *Nature (London)*, **400**, 342–344.
Miao, J. W., Hodgson, K. O., Ishikawa, T., Larabell, C. A., LeGros, M. A. & Nishino, Y. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 110–112.
Miao, J. W., Sayre, D. & Chapman, H. N. (1998). *J. Opt. Soc. Am.* **15**, 1662–1669.
Nellist, P. D. & Rodenburg, J. M. (1998). *Acta Cryst.* A**54**, 49–60.
Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature (London)*, **406**, 752–757.
Nugent, K. A., Gureyev, T. E., Cookson, D. F., Paganin, D. & Barnea, Z. (1996). *Phys. Rev. Lett*. **77**, 2961–2964.
Nugent, K. A., Peele, A. G., Chapman, H. N. & Mancuso, A. P. (2003). *Phys. Rev. Lett.* **91**, 203902.
Oszlányi, G. & Sütő, A. (2004). *Acta Cryst.* A**60**, 134–141.
Robinson, I. K., Vartanyants, I. A., Williams, G. J., Pfeifer, M. A. & Pitney, J. A. (2001). *Phys. Rev. Lett*. **87**, 195505.
Saleh, B. E. A. & Teich, M. C. (1991). *Fundamentals of Photonics.* New York: John Wiley and Sons.
Sneddon, I. N. (1951). *Fourier Transforms.* New York: McGraw-Hill.
Teague, M. R. (1983). *J. Opt. Soc. Am.* **73**, 1434–1441.