Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The oversampling phasing method

J. Miao,*† J. Kirz and D. Sayre

Department of Physics and Astronomy, State University of New York at Stony Brook, New York 11794, USA

† Mailing address: Stanford Synchrotron Radiation Laboratory, PO Box 4349 MS69, Stanford, California 94309-0210, USA.

Correspondence e-mail: miao@ssrl.slac.stanford.edu

Received 23 February 2000
Accepted 22 June 2000

Sampling the diffraction pattern of a finite specimen more finely than the Nyquist frequency (the inverse of the size of the diffracting specimen) corresponds to surrounding the electron density of the specimen with a no-density region. When the no-density region is bigger than the electron-density region, sufficient information is recorded so that the phase information can be retrieved from the oversampled diffraction pattern, at least in principle. By employing an iterative algorithm, the phase information from the oversampled diffraction pattern of a micrometre-sized test specimen has been successfully retrieved. This method is believed to be able to open a door for high-resolution three-dimensional structure determination of complex and non-crystalline biological specimens, *i.e.* whole cells and sub-micrometre molecular clusters and micrometre-sized protein crystals. With the possible appearance in the future of X-ray free-electron lasers, it may become possible to image single molecules by recording diffraction patterns before radiation damage manifests itself.

1. Introduction

When an object is illuminated by a plane wave, the diffracted wave in the far field, within the Born approximation, is the Fourier transform of the object. While the magnitude of the Fourier transform can be measured by a detector, the phase is lost. To recover the structure of the object, however, one has to know the phase information. This constitutes the well known phase problem. The phase problem is somewhat different for non-crystals and crystals. When the specimen is non-crystalline, the diffraction pattern is weak and continuous. This pattern can therefore be oversampled, *i.e.* sampled more finely than the Nyquist frequency. That oversampling a diffraction pattern could be used to retrieve the phase information in X-ray diffraction was suggested by Sayre (1991) on the basis of a similar method in optics (Bates, 1982). Historically, the oversampling method bore some relationship with solvent flattening (see, for example, Wang, 1985) and the use of non-crystallographic symmetry for phase determination (see, for example, Crowther, 1969). Recently, we proposed a theory to explain the oversampling method. We showed that sampling a diffraction pattern more finely than the Nyquist frequency corresponds to surrounding the electron density of the specimen with a no-density region. The finer the sampling, the bigger the no-density region. When the no-density region is bigger than the electron-density region, sufficient information is recorded so that at least in principle the phase can be retrieved from an oversampled diffraction pattern (Miao *et al.*, 1998). When the specimen is a crystal, the diffracted pattern is

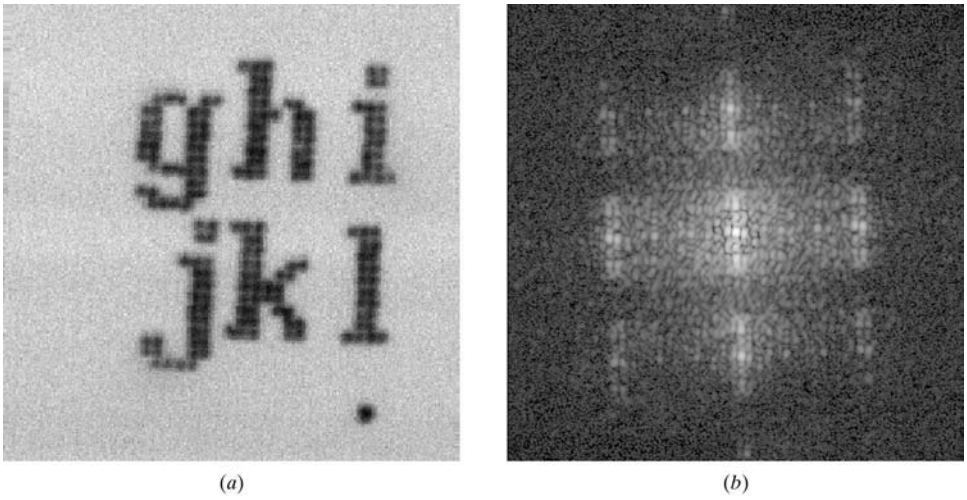

Figure 1
(*a*) A scanning transmission X-ray microscope image of the specimen. (*b*) A diffraction pattern of the specimen.

confined to discrete Bragg peaks owing to the constructive interference from many unit cells. These Bragg peaks are at the Nyquist frequency corresponding to a unit cell. This may explain why from Bragg peaks alone and without any *ab initio* information, the phase is not unique. If both the Bragg peaks and the intensity between the Bragg peaks, *i.e.* an oversampled diffraction pattern, can be measured from a finite crystal, the phase can be retrieved. For large finite crystals the intensity between Bragg peaks may be too weak to be measured. When the crystal is small, such as a micrometre-sized protein crystal, the intensity between Bragg peaks is no longer negligible. This extension of the oversampling method from non-crystalline specimens to small crystals is suggested and explored elsewhere (Miao & Sayre, 2000). In this method, crystal defects, if present, are accurately imaged.

2. Methods and results

Oversampling a diffraction pattern alone, however, cannot uniquely determine the phase since one cannot distinguish the correct phase and its conjugate from the diffraction pattern only. This twofold ambiguity can be eliminated by using positivity constraints. When the energy of the incident X-rays is high and away from the absorption edges, the electron density of the specimen is mostly real and positive. When the energy of incident X-rays is low, the electron density is complex, but the real and imaginary parts are both positive. These positivity constraints on the electron density can thus be used for phase determination. To obtain the phase information from an oversampled diffraction pattern, we have developed an iterative algorithm (Miao *et al.*, 1998; Miao & Sayre, 2000) by modifying that of Fienup (1982). Each iteration consists of the following four steps.

(i) By combining the measured magnitude of the Fourier transform and a guessed phase set, a new Fourier transform is assembled. (For the initial iteration, a random phase set was used.)

(ii) We then apply the inverse fast Fourier transform (FFT) to the assembled Fourier transform and obtain a new electron density.

(iii) Based on the level of oversampling of the diffraction pattern, a finite support is defined in real space to separate the no-density and the finite electron-density region. Owing to the difficulty of experimentally determining the exact envelope of the specimen, we choose the finite support to be bigger than the envelope of the specimen. We then enforce two kinds of constraints in real space. For the electron density outside the finite support, we force it close to zero. For the electron density inside, we enforce the positivity constraint and hence obtain a new electron density.

(iv) By applying FFT to the new electron density, we obtain a new Fourier transform and adopt its phase set. We then restore the phase of the central pixel to zero for a real electron density and obtain a new phase set.

After a few hundreds to thousands of iterations, convergence is usually complete.

By employing this algorithm, we have retrieved the phase information from an experimental oversampled diffraction pattern (Miao *et al.*, 1999). The diffraction pattern was recorded from a set of six letters with overall size about 5 μm. The letters, deposited on a silicon nitride membrane, were made up of gold dots each about 1000 Å in diameter and 800 Å in thickness. Fig. 1(*a*) shows a scanning transmission X-ray microscope (Kirz *et al.*, 1995) image of the specimen. The specimen was illuminated by a small and clean synchrotron X-ray beam ($\lambda = 17$ Å). The detector, a back-thinned and liquid-nitrogen-cooled CCD with 512 × 512 pixels and a 24 × 24 μm pixel size, was placed about 25 cm downstream of the specimen. To eliminate any unwanted scattering from the air, both the specimen and the detector were mounted in a vacuum with a pressure of 10^{-5}–10^{-6} Torr (1 Torr $\simeq$ 133.33 Pa). Fig. 1(*b*) shows an experimental diffraction pattern from the specimen in Fig. 1(*a*). In this figure a 19-pixel radius circular area at the center, a part of the pattern lost owing to a beam stop to block the direct beam, was filled by a patch derived from the magnitude of the Fourier transform calculated from Fig. 1(*a*). Although the patch occupied less than 0.5% of the whole diffraction pattern area, it was critical to the convergence of our algorithm.[1] We then input this diffraction

[1] We have found that the quality of the reconstruction is very sensitive to the accuracy of the data in the central patch. However, we have not performed any quantitative study of their relationship. In the long run, we plan to circumvent the problem by (i) recording the experimental diffraction pattern with the central patch as small as possible and (ii) enlarging our algorithm to allow it to iteratively fill in a small amount of data at the center.

pattern into our algorithm by employing a random initial phase set and a 7.5 × 7.5 μm square finite support. Figs. 2(*a*), 2(*b*), 2(*c*), 2(*d*) and 2(*e*) show the reconstructed density after 0, 100, 200, 300 and 400 iterations, respectively. One may notice that the density rotated 180° between Figs. 2(*b*) and 2(*c*), which was a consequence of the ambiguity of the correct phase and its conjugate. Fig. 2(*f*) shows the convergence of the algorithm where the error function was defined as the ratio of the total electron density outside the finite support to that of inside. The algorithm quickly pushed the electron density inside the finite support after 50 iterations, but took more iterations to find the correct phase set. After 400 iterations, a well reconstructed density (Fig. 3) was obtained. Fig. 3, evaluated more finely in the Fourier sum than Fig. 2(*e*) for display purposes, is consistent with the resolution limit, ∼600 Å, set by the angular extent of the CCD detector. The computing time for 400 iterations was about 15 min on a 450 MHz Pentium II workstation. We performed a few more reconstructions for the same diffraction pattern with different initial phase sets and found that the number of iterations required for convergence was somewhat different each time.

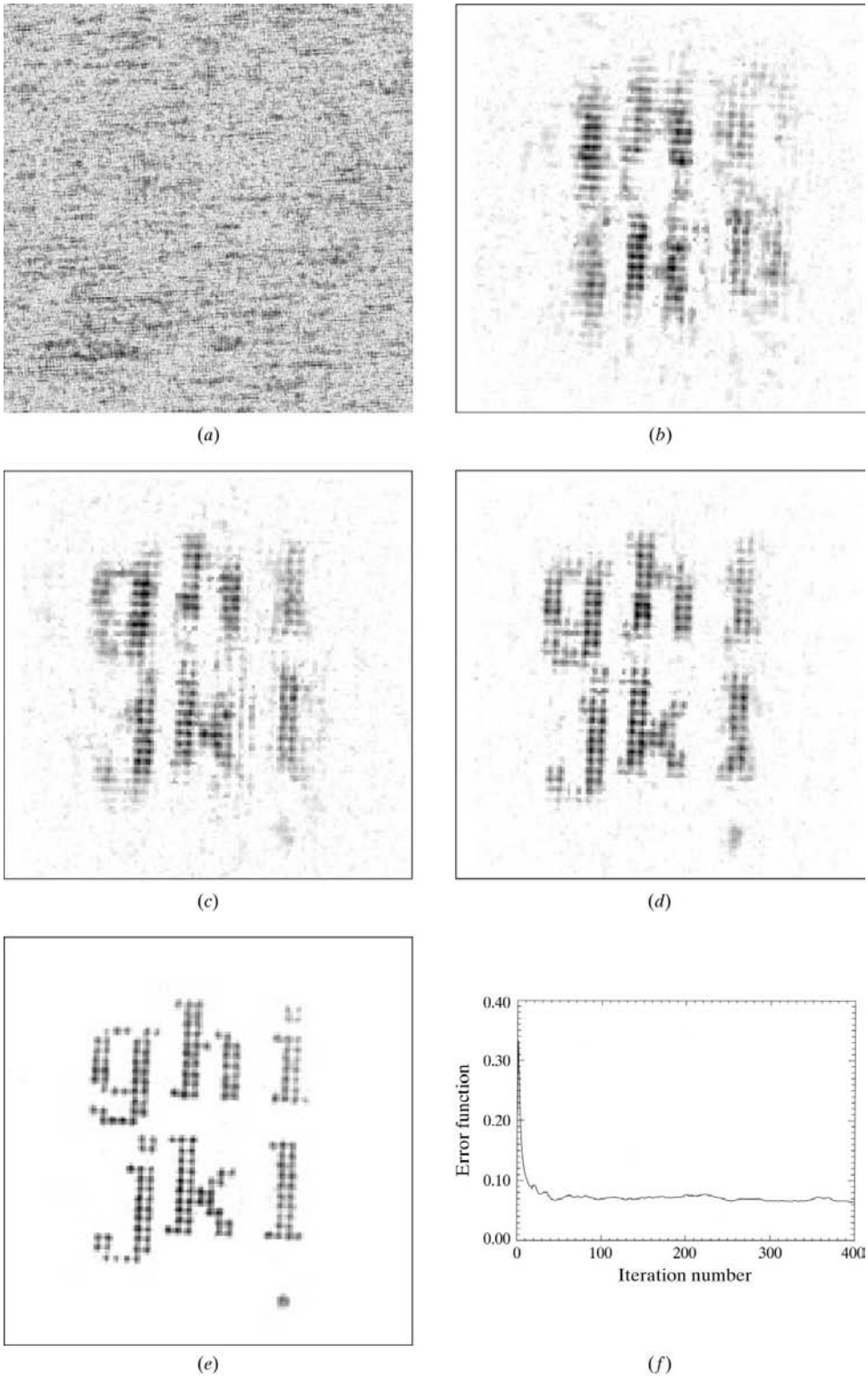

Figure 2
A reconstruction from the diffraction pattern of Fig. 1(*b*). (*a*) The initial input. (*b*) After 100 iterations. (*c*) After 200 iterations. (*d*) After 300 iterations. (*e*) After 400 iterations. (*f*) The convergence of the reconstruction.

3. Conclusions

We believe that this method can open a door for high-resolution three-dimensional structure determination of biological specimens such as whole cells and sub-micrometre molecule clusters. The only limitation to the resolution is radiation damage, which can be mitigated by using cryogenic techniques. Experimental studies (Maser *et al.*, 1998; Schneider & Niemann, 1998) show that biological specimens at liquid-nitrogen temperature can tolerate a radiation dose of up to 10^{10} Gy dosage without observable morphological damage. This method may also be extended to determine the structure of micrometre-sized protein crystals if both the Bragg peaks and the intensity between the Bragg peaks can be measured. In the long run, it may become possible to image single molecules by combining a fourth-generation X-ray source, such as a free-

Figure 3
A well reconstructed density after 400 iterations. (The scale bar corresponds to 1 μm.)

electron laser, with the oversampling phasing method. The challenge of this very intriguing goal is whether an atomic or nearly atomic resolution diffraction pattern can be recorded from single molecules before the structure is destroyed by radiation damage. The oversampling phasing method thus may shift the difficulty of growing crystals to overcoming the radiation-damage problem (Johnson & Blundell, 1999).

The idea of applying X-ray diffraction to three-dimensional structure determination of no-crystalline specimen was first suggested by Sayre (1980). The decision to try oversampling as a phasing technique was arrived at in a conversation in the late 1980s with G. Bricogne. We are grateful to P. Charalambous, Kings College, for fabrication of the test specimens. We thank C. Jacobsen for his help and advice with the numerical reconstruction, C. Jacobsen and M. Howells for use of the apparatus in which the exposures were made and S. Wirick for her assistance in the data acquisition. One of us (JM) thanks S. Doniach and K. Hodgson, Stanford Synchrotron Radiation Laboratory, for stimulating discussions. The experiment was performed at beamline X1A of the National Synchrotron Light Source, which is supported by the Department of Energy. This work is supported in part by Grant No. DE-FG0289ER60858 from the Department of Energy.

References

Bates, R. H. T. (1982). *Optik (Stuttgart)*, **61**, 247–262.
Crowther, R. A. (1969). *Acta Cryst.* B**25**, 2571–2580.
Fienup, J. R. (1982). *Appl. Opt.* **21**, 2758–2769.
Johnson, L. N. & Blundell, L. (1999). *J. Synchrotron Rad.* **6**, 813–815.
Kirz, J., Jacobsen, C. & Howells, M. (1995). *Quart. Rev. Biophys.* **28**, 33–130.
Maser, J., Jacobsen, C., Kirz, J., Osanna, A., Spector, S., Wang, S. & Warnking, J. (1998). *X-ray Microscopy and Spectromicroscopy*, edited by J. Thieme, G. Schmahl, D. Rudolph & E. Umbach, pp. 35–44. Berlin: Springer-Verlag.
Miao, J., Charalambous, C., Kirz, J. & Sayre, D. (1999). *Nature (London)*, **400**, 342–344.
Miao, J. & Sayre, D. (2000). *Acta Cryst.* A**56**. In the press.
Miao, J., Sayre, D. & Chapman, H. N. (1998). *J. Opt. Soc. Am. A*, **15**, 1662–1669.
Sayre, D. (1980). *Imaging Processes and Coherence in Physics. Springer Lecture Notes in Physics*, Vol. 112, edited by M. Schlenker, pp. 229–235. Berlin: Springer.
Sayre, D. (1991). *Direct Methods of Solving Crystal Structure*, edited by H. Schenk. *NATO ASI Series B (Physics)*, Vol. 274, pp. 353–356. New York: Plenum Press.
Schneider, G. & Niemann, B. (1998). *X-ray Microscopy and Spectromicroscopy*, edited by J. Thieme, G. Schmahl, D. Rudolph & E. Umbach, pp. 25–34. Berlin: Springer-Verlag.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.