Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 9 November 2000
Accepted 28 May 2001

Diffraction of partially coherent X-rays and the crystallographic phase problem

Abraham Szöke

Department of Biochemistry, Biomedical Centre, Box 576, Uppsala University, S-75123 Uppsala, Sweden, and Lawrence Livermore National Laboratory, Livermore, CA 94551, USA. Correspondence e-mail: szoke1@llnl.gov

This paper proposes a new physical method for the partial solution of the crystallographic phase problem by illuminating the crystal with an X-ray beam of limited coherence. The diffraction spots broaden and, if the coherence length of the incident beam is small enough in all three dimensions, the diffraction pattern becomes continuous. Independent information about the structure is then available both at the Bragg angles and at angles that do not satisfy the Bragg conditions. Under certain conditions, the total information is sufficient to solve the crystal structure *ab initio*. Two prescriptions for producing X-ray beams with limited coherence are given.

1. Introduction

When monochromatic coherent X-rays are diffracted from ideal crystals, diffracted beams are observed at discrete angles that satisfy the Bragg condition. Their intensities can be measured but their phases cannot. Reflection amplitudes with different phases correspond to different structures. Therefore, in general, a crystal structure cannot be recovered uniquely from its diffraction pattern alone. This is the well known phase problem of crystallography. The phase problem was expressed very clearly by Sayre (1952) in terms of information theory: the intensities of the diffraction pattern of a crystal, by themselves, do not contain enough information to determine the electron density of the crystal uniquely. In this paper, I will show that a continuous diffraction pattern is produced if a crystal is illuminated by an X-ray beam with small coherence length. Independent information about the structure is therefore available both at the Bragg angles and at angles that do not satisfy the Bragg conditions. Under certain conditions, the total information is sufficient to solve the crystal structure *ab initio*.

The plan of the paper is as follows. I start with a brief survey of some of the existing methods to obtain and to use additional information for the solution of crystal structures. Still in this *Introduction*, I will show the plausibility of my proposal, by way of its analogy to some holographic techniques. As the thesis of the paper is far from obvious, I will try to give a fairly detailed theoretical background. In the main part of the paper, I start with a rather pedantic derivation of the formulae for the diffraction of partially coherent X-rays on crystals. Then I discuss two methods to produce incident X-ray beams with the required (lack of) coherence. As these two sections are rather unfamiliar to many readers, I will try to define clearly what is meant by coherence in space and time. I follow with a discussion of the 'solution of the phase problem' from three different points of view. I will try to clarify the connection between uniqueness of the solution of the crystal structure and adequate sampling of the diffraction pattern from a holographic point of view, and I will finish with some practical recovery methods. In the discussion section, I will recapitulate the main arguments and try to convince the reader that the main results of the paper can also be derived from a point of view that is more familiar to crystallographers.

The proposed new method will make a crystal give up more of its secrets. It is my conviction that reliable direct experimental information is the most important ingredient in advancing our understanding of difficult problems in crystallography.

1.1. Background

Over the years, crystallographers have developed 'physical' and 'non-physical' methods to solve the phase problem (Woolfson & Fan, 1995). Physical methods attempt to collect sufficient independent experimental information about a crystal, while non-physical methods utilize external or *a priori* information.[1] The most important physical methods are multiple isomorphous replacement (MIR) and multi-wavelength anomalous dispersion (MAD). The first is based on measuring the diffraction patterns of closely related

[1] Some other methods in crystallography do not fit neatly into the above categories. The most basic *a priori* information is the non-negativity of the electron density. In addition, one can use the relative flatness of the electron density in parts of the crystal occupied by disordered solvent and a partial knowledge of the molecule (molecular replacement). For the purposes of this paper, their discussion is not needed. I would like to mention, but not discuss in detail, three more physical methods. One is based on multiple-beam reflections that determine the absolute phases of triplet invariants (Weckert & Hümmer, 1997, and references). The other two are the happenstances of non-crystallographic symmetry and the crystallization of the same molecule in two different space groups. Both of the latter give (independent) information on the amplitudes of the molecular diffraction pattern in additional directions.

structures. The second utilizes the fact that, near its X-ray absorption edge, the scattering amplitude of an atom depends on the incident wavelength. The pioneering non-physical methods, known as direct methods, use atomicity together with non-negativity of the electron density. They are based on the fact that, at high resolution, there are many more reflections in the diffraction pattern of a crystal than the number of atoms whose positions we are trying to recover and that those atoms occupy only a fraction of the crystal volume. Therefore there is enough information to find the atom positions. The most useful and widespread non-physical method exploits chemical information: in most cases, the compositions of the molecules in the crystal are known and the crystallographer builds a molecular model that agrees with the known bond lengths and angles as well as with the measured diffraction intensities.

This paper will refer often to the connection between holography and X-ray crystallography (Tollin *et al.*, 1966; Szöke, 1993; Szöke *et al.*, 1997). In order to describe that connection clearly, I will assume now that part of the electron-density structure in the unit cell of a crystal is known and an X-ray diffraction experiment is carried out. The wave diffracted from the known part can then be viewed as a holographic reference wave and the wave diffracted from the unknown part of the unit cell can be viewed as the holographic object wave. The resultant diffraction pattern contains, in addition to the sum of the intensities of the reference wave and the object wave, an interference term that is analogous to a hologram. Because the incident X-ray beam is coherent over many unit cells, the waves scattered from different unit cells also interfere with one another. The result is a sampling of the diffraction pattern at discrete diffraction angles that satisfy the Bragg condition. The sampling is sparse enough that the unknown scatterer cannot be reconstructed uniquely from the hologram (Sayre, 1952; Bricogne, 1992). This is, obviously, a restatement of the phase problem of X-ray crystallography in terms of holography.[2]

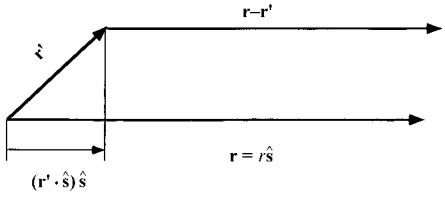

Figure 1
Scattering geometry.

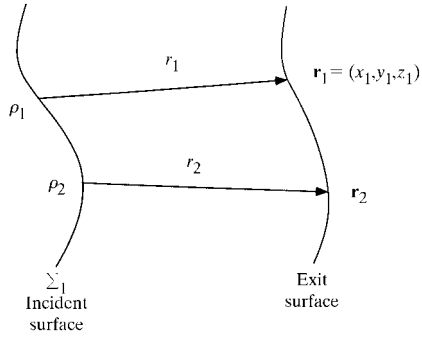

Figure 2
Geometry for propagation of mutual coherence.

1.2. Methods to obtain continuous diffraction patterns

Let us imagine a collection of well aligned molecules that are disordered in their positions. One possible example would be a molecular gas whose molecules are all oriented the same way (possibly by the application of an electric or magnetic field). A more practical example would be a crystal in which each position is occupied by one or the other of two similar molecules at random (Szöke, 1999). If a monochromatic X-ray or electron beam is incident on such a crystal, we obtain a discrete (Bragg) diffraction pattern and a continuous (diffuse) one. The Bragg diffraction is that of the average crystal and the diffuse pattern is that of a single 'difference' molecule. The latter can be measured at any diffraction angle: there are no Bragg conditions to satisfy and the interference terms among different molecules average out to zero (Szöke, 1986, 1997, 1999).[3] Earlier, Saldin & de Andres (1990) proposed that a diffuse diffraction pattern can be obtained when low-energy electrons illuminate a crystalline surface that has relatively few atoms adsorbed on it. They pointed out that the resulting diffuse low-energy electron diffraction pattern is a hologram and, indeed, it has been observed experimentally and reconstructed successfully by holographic methods (Saldin, 1997).

The main purpose of this paper is to show that a similar situation arises in ordinary crystals when the incident X-ray or electron beam is only partially coherent. The diffraction

[2] The origins of this paper stem from my previous work on holographic microscopy with an internal source of radiation (Szöke, 1986), where some of the important conclusions of this paper were also presented without detailed proof. I proposed that X-rays or photoelectrons emitted by an excited atom are analogous to a reference wave in holography. If the emitting atom is surrounded by other atoms, as in a crystal, part of the reference wave scatters from the surrounding atoms. The scattered waves are analogous to the object wave in holography. When these two waves arrive at a distant screen, their interference term is analogous to a hologram (Gábor, 1948, 1949). Moreover, in a crystal there are many identical emitters in different unit cells surrounded by identical environments, each one producing a similar hologram. The waves emitted by different excited atoms in the crystal are incoherent. The result is that the hologram intensity recorded on the distant screen is the sum of the intensities produced by each emitter alone. This results from a statistical cancellation of all interference terms among the waves emitted by different atoms. Such holograms have been observed both with electrons (see *e.g.* Saldin, 1997) and with X-rays (Tegze & Faigel, 1996; Gog *et al.*, 1996). Several algorithms have been used successfully to recover the three-dimensional structure of the emitting atoms' environment. I should mention some of the important properties of internal-source holograms that enable atomic resolution images in three dimensions. The most important property is that the wavelength of the emitted radiation is short and it is feasible to record the hologram on the full sphere. Second, the emitter and the scatterer are at atomic distances from each other. This provides stability, magnification, a much relaxed need for coherence of the source and no resolution limitation on the recording medium. (All the above are problematic in 'conventional' holography.) Third, the hologram records only the immediate vicinity of the emitter. Therefore, it is insensitive to long-range order in the crystal. Its main disadvantage is the very low contrast (signal level) of the hologram. There are two important differences between holographic microscopy and X-ray crystallography. First, the interference terms among the individual holograms average to zero. Second, there is enough information in a single hologram to reconstruct the scattering object (albeit possibly with some ambiguity).

[3] In 1975, in collaboration with A. Hawryluk, I performed model-based experiments that verify this assertion, but they were never published.

pattern produced by each coherence volume is incoherent with the diffraction pattern produced by any other coherence volume. Therefore, the diffraction pattern of the whole crystal is their incoherent sum. The angular distribution of the diffraction pattern is that of a single coherence volume, but its total intensity is characteristic of the whole crystal.

In summary, I know of three different ways to produce a continuous diffraction pattern of molecules. The first type of diffraction pattern is the X-ray hologram produced by independently emitting atoms in a crystal (as discussed in the footnote to §1.1). The second one is produced by X-rays incident on a translationally disordered ensemble of well oriented molecules. The third one is produced when partially coherent X-rays irradiate a well ordered crystal. Each one of them is the incoherent sum of identical interference patterns produced within a small neighborhood, and therefore the damage is distributed over all the molecules. Each one of them has enough information to recover the scatterer uniquely (within some limits) and each one can be reconstructed, or 'solved' by its analogy to holography. The first two scenarios have been discussed previously (Szöke, 1986, 1997, 1999). In this paper, I discuss the third scenario.

2. A simple explanation

In view of the rather weighty mathematics that follow, I will try to give now some simple explanations and then point out some important distinctions between present-day practice and the proposed new method.

Suppose we could concentrate our X-ray beam so as to illuminate a single unit cell of a crystal. The diffraction pattern observed would obviously be that of a single unit cell: a non-periodic electron density (Miao *et al.*, 1999; Neutze *et al.*, 2000). Now, let us irradiate a different unit cell with another beam from a different source. As the two sources are completely independent, the intensity of the two diffraction patterns is the sum of the intensities of the individual diffraction patterns. The essence of my proposition is the following: if the two unit cells are illuminated by the same beam, even though the amplitudes and phases of this beam vary independently in the two unit cells, the same thing happens.

In contrast, if a monochromatic X-ray beam is focused (mildly) onto a crystal, a Bragg diffraction pattern is observed. Note that the focused X-ray beam has a transverse coherence length of $\lambda/(\mathrm{NA})$, where (NA), called the numerical aperture, is roughly the convergence angle of the incident X-ray beam (in radians). The transverse coherence length can often be less than the transverse dimensions of the unit cell. Alternatively, suppose now that we shine a parallel X-ray beam of wide spectral width onto a crystal. I just described a Laue diffraction experiment. The longitudinal coherence length of the incident beam is the X-ray wavelength λ divided by the fractional bandwidth, $\Delta\lambda/\lambda$, of the incident radiation. In many Laue diffraction experiments, the longitudinal coherence of the incident radiation is much shorter than the unit cell. It is clear that we need both the short parallel coherence length of the focused beam and the short longitudinal coherence length of the non-monochromatic beam to obtain what we want!

3. Diffraction of partially coherent electromagnetic waves

In this section, we will derive the formulae for the scattering of partially coherent X-rays on crystals. The central results are equations (18) and (22). They establish the theoretical basis of our claims.

It has been realized for many years that some properties of scattered radiation are influenced by spectral and coherence properties of the incident electromagnetic waves (Mandel, 1969; van Kampen, 1969). We will follow the paper by Wolf & Foley (1989), summarized by Mandel & Wolf (1995).

X-ray scattering in a crystal occurs because its electron density varies in space and time. The incident radiation induces a current in the scattering material or, equivalently, a change in its polarization density. As the incident X-ray field is weak, static and dynamic properties of the crystalline scatterer are not influenced by it. This is called the linear-response approximation (Fetter & Walecka, 1971). Under these conditions, the response of the material, as well as the amplitude of the scattered wave, can be derived from first principles of quantum electrodynamics.

3.1. Scattering theory

The scattering process can be described equally well by classical electrodynamics with a phenomenological linear susceptibility. We will assume that the material is non-magnetic and non-conducting. The polarization induced in the material will be denoted by $\mathbf{P}(\mathbf{r}, t)$, where $\mathbf{r}$ is the coordinate of the scatterer and t is the time of observation. It is proportional to the electric field of the electromagnetic radiation at an earlier time, $\mathbf{E}(\mathbf{r}, t-\tau)$. Their ratio is called the linear susceptibility $\eta(\mathbf{r}, t; \tau)$.[4] The explicit time dependence of the susceptibility, denoted here by t, signifies the time dependence of the crystal structure as a result of its thermal motion. The susceptibility also depends on the time delay, $\tau \geq 0$, between the action of the electric field and the observation time t of the polarization caused by it. The constitutive relation of the material can then be written as (Wolf & Foley, 1989)

$$\mathbf{P}(\mathbf{r}, t) = \frac{1}{2\pi}\int_0^\infty \mathrm{d}\tau\, \eta(\mathbf{r}, t; \tau)\, \mathbf{E}(\mathbf{r}, t-\tau). \qquad (1)$$

The susceptibility of crystals is generally a second-order tensor for visible light. For X-rays, a scalar η is sufficient (although it is advantageous to allow it to be a complex number). The two separate time dependencies of $\eta(\mathbf{r}, t; \tau)$ can be clarified by applying to it a double Fourier transformation:

[4] In the first 18 equations we use real notation.

$$\hat{\eta}(\mathbf{r}, \mathbf{\Omega}; \omega_0) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \mathrm{d}t \exp(i\mathbf{\Omega}t) \int_{0}^{\infty} \mathrm{d}\tau\, \eta(\mathbf{r}, t; \tau) \exp(i\omega_0\tau). \tag{2}$$

It can be shown, using (2), that the time dependence of the crystal structure produces scattered waves at frequencies $\omega_0 \pm \Omega$ around the incident frequency ω_0: they are called Brillouin and Raman components (Berne & Pecora, 1976). It is also apparent that the time delay τ is important only when there is a strong ω_0 dependence of η within the frequency spectrum of the incident light. This can occur when the incident electric field frequency is close to a material resonance (*e.g.* an ionization edge).

If equation (1) is substituted into Maxwell's equations, the wave equation for the scattered field can be obtained. We refer to the books of Born & Wolf (1980, §2.2.2) or Jackson (1975, §9.7). Equation 9.100 of Jackson can be written as

$$\nabla^2 \mathbf{D}(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2 \mathbf{D}(\mathbf{r}, t)}{\partial t^2} = -4\pi \nabla \times [\nabla \times \mathbf{P}(\mathbf{r}, t)], \tag{3}$$

where $\mathbf{D}$ is the displacement vector, related to $\mathbf{E}$ by

$$\mathbf{D}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) + 4\pi \mathbf{P}(\mathbf{r}, t). \tag{4}$$

We introduce now the Hertz vector, $\mathbf{\Pi}$, which is related to $\mathbf{D}$ by

$$\mathbf{D}(\mathbf{r}, t) = \nabla \times [\nabla \times \mathbf{\Pi}(\mathbf{r}, t)]. \tag{5}$$

A formal solution of (3) can be obtained as the volume integral of the retarded Green's function

$$\mathbf{\Pi}(\mathbf{r}, t) = \int_V \frac{\mathbf{P}(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/c)}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r}', \tag{6}$$

where the three-dimensional integral $\mathrm{d}\mathbf{r}'$ is over the scattering volume, V. In the following, we will distinguish between the scattering region by using $\mathbf{r}'$ for its coordinates and $\mathbf{r}$ for the coordinates of the region of observation.

The set of integral equations (1)–(6) is similar to the Lippmann–Schwinger equations in quantum scattering theory (Newton, 1982). Their solutions have to satisfy scattering boundary conditions. Namely, at $t \to -\infty$ there is only an incident wave and at large distances from the scatterer the solution consists of the original incident wave and a scattered wave that has only outgoing components.

3.2. Scattering of partially coherent X-rays on crystals

We now develop an expression for the scattered intensity, following Wolf & Foley (1989), but using a time-dependent formalism. First, we assume that the spectrum of the incident (and scattered) X-rays is narrow. The incident field, $\mathbf{E}^{(i)}(\mathbf{r}', t)$, is then characterized by a center frequency ω_0 and an associated wavevector of magnitude $k_0 = \omega_0/c$.

Second, we assume that the incident spectrum is far enough from any material resonance that the time delay of the material response can be neglected. Then (1) simplifies to

$$\mathbf{P}(\mathbf{r}', t) = \eta(\mathbf{r}', t; 0)\, \mathbf{E}(\mathbf{r}', t). \tag{7}$$

Third, we assume that the scattering is weak. Then the equations can be solved in the first Born approximation. The field in (7) that induces the polarization is the incident field $\mathbf{E}^{(i)}(\mathbf{r}', t)$. If the substitution of the incident field for $\mathbf{E}(\mathbf{r}', t)$ is not made in (6), the full dynamical theory of Laue follows. The Hertz vector from (6) gives the scattered field

$$\mathbf{\Pi}(\mathbf{r}, t) = \int_V \frac{\eta(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/c; 0)\, \mathbf{E}^{(i)}(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/c)}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r}'. \tag{8}$$

Fourth, we assume that the size of the scatterer is much smaller than the distance to the screen where we make our observations: $|\mathbf{r}'| \ll |\mathbf{r}|$. This is the 'far field' approximation. We remind the reader that primed coordinates are in the scattering volume, V, and unprimed coordinates denote the screen, where the scattered field is observed. As seen in Fig. 1, the distance can be approximated as $|\mathbf{r} - \mathbf{r}'| \simeq r - \mathbf{r}' \cdot \hat{\mathbf{s}}$; here $\hat{\mathbf{s}}$ is a unit vector in the direction of the scattered wave $\mathbf{r}$, $r = |\mathbf{r}|$ and $\mathbf{r}' \cdot \hat{\mathbf{s}}$ denotes the scalar product of the two vectors. (Therefore, $\mathbf{r} = r\hat{\mathbf{s}}$.) In the denominator, the cruder approximation $|\mathbf{r} - \mathbf{r}'| \simeq r$ can be used. Furthermore, we define

$$t' = t - r/c, \tag{9}$$

the average retarded time at the scatterer. This gives

$$\mathbf{\Pi}(\mathbf{r}, t) = \frac{1}{r} \int_V \eta(\mathbf{r}', t' + \mathbf{r}' \cdot \hat{\mathbf{s}}/c; 0)\, \mathbf{E}^{(i)}(\mathbf{r}', t' + \mathbf{r}' \cdot \hat{\mathbf{s}}/c) \, \mathrm{d}\mathbf{r}'. \tag{10}$$

The scattered field can be calculated from (5):

$$\mathbf{D}^{(s)}(\mathbf{r}, t) = \nabla \times \left[\nabla \times \frac{1}{r} \int_V \eta(\mathbf{r}', t' + \mathbf{r}' \cdot \hat{\mathbf{s}}/c; 0) \times \mathbf{E}^{(i)}(\mathbf{r}', t' + \mathbf{r}' \cdot \hat{\mathbf{s}}/c) \, \mathrm{d}\mathbf{r}' \right], \tag{11}$$

where $\nabla \times [\nabla \times \cdot]$ operates on $\mathbf{r}$.

Fifth, we assume that near the screen there is no dielectric medium; therefore, $\mathbf{D}^{(s)}(\mathbf{r}, t) = \mathbf{E}^{(s)}(\mathbf{r}, t)$. From the narrow-band and far-field approximations, it follows that

$$\nabla \times [\nabla \times \cdot] \simeq k_0^2\, \hat{\mathbf{s}} \times [\hat{\mathbf{s}} \times \cdot],$$

where k_0 is the magnitude of the wavevector at the center frequency of the incident electromagnetic waves: $k_0 = \omega_0/c$. In summary,

$$\mathbf{E}^{(s)}(\mathbf{r}, t) = \frac{k_0^2}{r} \int_V \eta(\mathbf{r}', t' + \mathbf{r}' \cdot \hat{\mathbf{s}}/c; 0)\, \hat{\mathbf{s}} \times [\hat{\mathbf{s}} \times \mathbf{E}^{(i)}(\mathbf{r}', t' + \mathbf{r}' \cdot \hat{\mathbf{s}}/c) \, \mathrm{d}\mathbf{r}']. \tag{12}$$

Our sixth assumption is that we measure the total intensity of the (narrow-band) scattered waves at the screen and

not its spectral distribution.[5] The time-dependent intensity is given by

$$\begin{aligned} I^{(s)}(\mathbf{r},t) &= (c/4\pi)|\mathbf{E}^{(s)}(\mathbf{r},t)|^2 \\ &= \frac{c}{4\pi}\frac{k_0^4}{r^2}\int_V \mathrm{d}\mathbf{r}' \int_V \mathrm{d}\mathbf{r}''\big(\eta(\mathbf{r}', t'+\mathbf{r}'\cdot\hat{\mathbf{s}}/c;0) \\ &\quad \times \eta(\mathbf{r}'', t'+\mathbf{r}''\cdot\hat{\mathbf{s}}/c;0)\{\hat{\mathbf{s}}\times[\hat{\mathbf{s}}\times\mathbf{E}^{(i)}(\mathbf{r}', t'+\mathbf{r}'\cdot\hat{\mathbf{s}}/c)]\} \\ &\quad \cdot\{\hat{\mathbf{s}}\times[\hat{\mathbf{s}}\times\mathbf{E}^{(i)}(\mathbf{r}'', t'+\mathbf{r}''\cdot\hat{\mathbf{s}}/c)]\}\big). \end{aligned} \tag{13}$$

The dependence of this expression on the polarization of the incident radiation can be simplified by repeated use of the vector identity

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = (\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D}) - (\mathbf{A}\cdot\mathbf{D})(\mathbf{B}\cdot\mathbf{C})$$

among the arbitrary three-dimensional vector, **A**, **B**, **C** and **D**. The relations are applied to combinations of the vectors $\hat{\mathbf{s}}$ and $\mathbf{E}^{(i)}$. The notation can be further simplified by introducing the Cartesian components s_n and $E_n^{(i)}$ of $\hat{\mathbf{s}}$ and $\mathbf{E}^{(i)}$, respectively:

$$\begin{aligned} &I^{(s)}(\mathbf{r},t) \\ &= \frac{c}{4\pi}\frac{k_0^4}{r^2}\sum_{m=1}^{3}\sum_{n=1}^{3}\Bigg\{(\delta_m n - s_m s_n) \\ &\quad \times \int_V \mathrm{d}\mathbf{r}' \int_V \mathrm{d}\mathbf{r}''\Big[\eta(\mathbf{r}', t'+\mathbf{r}'\cdot\hat{\mathbf{s}}/c;0)\,\eta(\mathbf{r}'', t'+\mathbf{r}''\cdot\hat{\mathbf{s}}/c;0) \\ &\quad \times E_m^{(i)}(\mathbf{r}', t'+\mathbf{r}'\cdot\hat{\mathbf{s}}/c)\,E_n^{(i)}(\mathbf{r}'', t'+\mathbf{r}''\cdot\hat{\mathbf{s}}/c)\Big]\Bigg\}. \end{aligned} \tag{14}$$

Our seventh assumption is that the scattered intensity is integrated over times that are long with respect to the correlation time of the thermal motion in the crystal and the correlation time of the incident radiation. Our eighth assumption is that the fluctuations in the thermal motion of the crystal and those in the X-ray emission are uncorrelated. These allow us to define the time-averaged intensity and to decompose the time average of the integrand in (14) into the product of those of the susceptibility and of the incident waves:

$$\begin{aligned} \langle I^{(s)}(\mathbf{r},t)\rangle &= \lim_{T\to\infty}\frac{1}{T}\int_{t-T}^{t} I^{(s)}(\mathbf{r},\tau)\,\mathrm{d}\tau \\ &= \frac{c}{4\pi}\frac{k_0^4}{r^2}\sum_{m=1}^{3}\sum_{n=1}^{3}\Bigg\{(\delta_{mn} - s_m s_n) \\ &\quad \times \int_V \mathrm{d}\mathbf{r}' \int_V \mathrm{d}\mathbf{r}''\big[\langle\eta(\mathbf{r}', t'+\mathbf{r}'\cdot\hat{\mathbf{s}}/c;0)\,\eta(\mathbf{r}'', t'+\mathbf{r}''\cdot\hat{\mathbf{s}}/c;0)\rangle \\ &\quad \times \langle E_m^{(i)}(\mathbf{r}', t'+\mathbf{r}'\cdot\hat{\mathbf{s}}/c)E_n^{(i)}(\mathbf{r}'', t'+\mathbf{r}''\cdot\hat{\mathbf{s}}/c)\rangle\big]\Bigg\}, \end{aligned} \tag{15}$$

where the first equality defines the time average $\langle\cdot\rangle$.

[5] In dynamic light-scattering spectroscopy, the spectral distribution of the scattered light carries important information on the internal motion of the scattering material (Berne & Pecora, 1976). We used a time-dependent formalism in order to express the integrated intensity conveniently. The formalism used by Wolf & Foley (1989) is better suited to calculate the spectrum of the scattered light.

Our ninth assumption is that both the thermal motion and the incident radiation are stationary random processes, at least in the wide sense (Papoulis, 1991). Accordingly, the time averages depend only on the difference between the two time arguments in the expressions.

The two factors in the double integral in equation (15) have well established names and meanings. We will first introduce the notation and discuss them in detail in §3.3. The first factor is called the (generalized) Van Hove correlation function for the susceptibility of the crystal:

$$G(\mathbf{r}';\mathbf{r}'';T) = \langle\eta(\mathbf{r}',t;0)\,\eta(\mathbf{r}'',t+T;0)\rangle. \tag{16}$$

The other factor, $\langle E_m^{(i)}(\cdot)E_n^{(i)}(\cdot)\rangle$, is the time-averaged correlation function of the incident electromagnetic field, measured at two different points and two different times. It is defined in books by Born & Wolf (1980, ch. X) and Goodman (1985, chs. 5 and 7). It is called the mutual-coherence function. We use the notation

$$\Gamma_{m,n}^{(i)}(\mathbf{r}';\mathbf{r}'';T) = \langle E_m^{(i)}(\mathbf{r}',t)\,E_n^{(i)}(\mathbf{r}'',t+T)\rangle. \tag{17}$$

Using these definitions, (15) can be written as

$$\begin{aligned} \langle I^{(s)}(\mathbf{r},t)\rangle &= \frac{c}{4\pi}\frac{k_0^4}{r^2}\sum_{m=1}^{3}\sum_{n=1}^{3}\Bigg[(\delta_{mn} - s_m s_n) \\ &\quad \times \int_V \mathrm{d}\mathbf{r}' \int_V \mathrm{d}\mathbf{r}''\,G(\mathbf{r}';\mathbf{r}'';(\mathbf{r}'-\mathbf{r}'')\cdot\hat{\mathbf{s}}/c) \\ &\quad \times \Gamma_{m,n}^{(i)}(\mathbf{r}';\mathbf{r}'';(\mathbf{r}'-\mathbf{r}'')\cdot\hat{\mathbf{s}}/c)\Bigg]. \end{aligned} \tag{18}$$

This is our most important result. In words, it says that the diffracted X-ray intensity is proportional to the time-averaged mutual correlation function of the crystal susceptibility multiplied by the mutual coherence function of the incident electromagnetic field.

3.3. Detailed interpretation

As the concept of mutual coherence is not entirely intuitive, I will discuss it further, bringing in some optical jargon. In interference experiments, two light beams are brought together and the resulting intensity is measured. If the light beams have the appropriate space–time relationship, such experiments can be viewed as a measurement of the correlation of the electromagnetic field amplitudes at two different points in space and time. If the electromagnetic fields are produced by a stationary random process, the time average of such a measurement is a constant. It measures the mutual coherence function of the electromagnetic field. It answers the question: to what extent does knowledge of the field at one point and one time have predictive power about the field at the other point and the other time? If the predictive power is high, the field is called coherent and, if it is low, the field is incoherent at those two points in space and time. All second-order interference phenomena depend only on the mutual coherence function of the field.

In order to gain some insight into the possible implications of (18), we will work out the consequences of an overly

simplified model for the incident beam. (A more realistic model, presented in Appendix *B*, reaches slightly different conclusions.) We assume that it is a randomly modulated plane wave, in the 'general' direction $\mathbf{k}^{(i)}$ and that its coherence falls off in all three spatial dimensions as a Gaussian, $\exp(-|\mathbf{r}' - \mathbf{r}''|^2/d^2)$, where $d \gg \lambda_i$ is some coherence length. We will then approximate the mutual coherence function of the incident beam, for any T, by[6]

$$\Gamma^{(i)}_{m,n}(\mathbf{r}'; \mathbf{r}''; T) \simeq E^{(i)}_m E^{(i)}_n \exp[i\mathbf{k}^{(i)} \cdot (\mathbf{r}' - \mathbf{r}'') - i\omega_i T] \exp(-|\mathbf{r}' - \mathbf{r}''|^2/d^2). \quad (19)$$

We assume, as usual, that the unit cell is much larger than the wavelength. Nevertheless, the size of the incident X-ray beam can be much larger than d; in that case, the incident X-ray beam is an incoherent superposition of many coherent, almost parallel, beamlets. By assuming a longitudinal coherence, we assumed that the spectral width of the beam is $\Delta\lambda_i/\lambda_i \simeq \lambda_i/d$.

Using $T = (\mathbf{r}' - \mathbf{r}'') \cdot \hat{\mathbf{s}}/c$ as in (18), elastic scattering, $\omega_i = \omega_o$, and, therefore, $\omega_o \hat{\mathbf{s}}/c = \mathbf{k}^{(s)}$, we obtain

$$\Gamma^{(i)}_{m,n}(\mathbf{r}'; \mathbf{r}''; (\mathbf{r}' - \mathbf{r}'') \cdot \hat{\mathbf{s}}/c) \simeq E^{(i)}_m E^{(i)}_n \exp[i(\mathbf{k}^{(i)} - \mathbf{k}^{(s)}) \cdot (\mathbf{r}' - \mathbf{r}'')] \exp(-|\mathbf{r}' - \mathbf{r}''|^2/d^2). \quad (20)$$

Let us now also ignore disorder and thermal motion in the crystal. Then,

$$G(\mathbf{r}'; \mathbf{r}''; T) \simeq G(\mathbf{r}'; \mathbf{r}''; 0) \simeq \rho(\mathbf{r}')\, \rho(\mathbf{r}''), \quad (21)$$

where we approximated the susceptibility $\eta(\mathbf{r}', t; 0)$ in (16) by the electron density $\rho(\mathbf{r}')$. Substituting (20) and (21) into (18), defining $\mathbf{R} = \mathbf{r}'' - \mathbf{r}'$ and using $\mathbf{R}$ and $\mathbf{r}'$ as integration variables, we obtain

$$\langle I^{(s)}(\mathbf{r}, t)\rangle = \frac{c}{4\pi}\frac{k_0^4}{r^2}\sum_{m=1}^{3}\sum_{n=1}^{3}\Bigg\{(\delta_{mn} - s_m s_n) \times \int_V \mathrm{d}\mathbf{r}' \int_V \mathrm{d}\mathbf{R}\, \rho(\mathbf{r}')\, \rho(\mathbf{r}' + \mathbf{R})\, E^{(i)}_m\, E^{(i)}_n \times \exp[-i(\mathbf{k}^{(s)} - k^{(i)}) \cdot \mathbf{R}] \exp(-|\mathbf{R}|^2/d^2)\Bigg\}. \quad (22)$$

In some sense, we can consider

$$\rho(\mathbf{r}')\, \rho(\mathbf{r}' + \mathbf{R}) \exp(-|\mathbf{R}|^2/d^2) \quad (23)$$

to be the Patterson function of a crystal of size d of 'Gaussian shape'.[7] The inner integral in (22) is the intensity of the diffraction pattern of that small crystal.

It is of interest to discuss the diffraction pattern of a crystal as we vary the coherence length d of the incident X-ray beam, while keeping the whole crystal illuminated. It will be shown in equation (24) that, for a fully coherent incident beam, equation (22) reduces to the usual formula for X-ray diffraction. Note, in particular, that the sharpness of the Bragg peaks is the consequence of the long-range order of the crystal or, in other words, of the existence of large Patterson vectors that are much longer than the unit cell. Suppose now that the crystal is composed of small perfect crystallites with slightly different orientations. As usual, we assume that the relative origins of the crystallites are random. If the coherence length of the incident X-rays is longer than the average size of the crystallites, we get a sum of the intensities of the diffraction patterns of the individual crystallites. If the coherence length becomes shorter than the sizes of the individual crystallites, the Bragg spots become somewhat broader and parts of the crystallites contribute to the diffraction incoherently.

If we make the coherence length d even shorter, only the short Patterson vectors will contribute to the diffraction pattern. As the coherence length of the incident radiation becomes smaller and smaller, the diffraction peaks broaden and, when d is of the size of a single unit cell, the diffraction is continuous: the Bragg conditions do not have to be satisfied in order to obtain appreciable diffraction intensity.

If the coherence length is much shorter than the unit cell (or the size of the molecule), we lose the Patterson vectors that give information about the relative positions of parts of the molecule. Clearly, this is not desirable and, therefore, there is an optimal coherence length for obtaining maximal information.

Note that the outer integral in (22) is over the whole illuminated area of the crystal. Therefore, if the crystal is illuminated by an X-ray beam of limited coherence, the total number of diffracted photons is the same as from a coherent incident beam, but, as indicated above, the diffracted intensity is not concentrated into Bragg peaks. This point cannot be overemphasized: the total integrated intensity of the diffraction pattern is independent of the coherence of the incident beam!

As promised in the *Introduction*, equation (22) shows the analogy between the diffraction of partially coherent X-rays from crystals and the diffraction of a coherent X-ray beam from an orientationally ordered but translationally disordered collection of molecules.

A final note for crystallographers who are not familiar with equation (18). I will show now that, for a coherent incident field, (18) reduces to the usual expressions for X-ray diffraction. Let us set $\exp(-|\mathbf{R}|^2/d^2) = 1$ in equation (22). This results in the well known expression for the diffraction intensity of an infinite crystal in the kinematic approximation (James, 1982; Jagodzinski & Frey, 1992):

$$\langle I^{(s)}(\mathbf{r}, t)\rangle_{\text{coherent}} = \frac{c}{4\pi}\frac{k_0^4}{r^2}\sum_{m=1}^{3}\sum_{n=1}^{3}\Bigg\{(\delta_{mn} - s_m s_n) \times \int_V \mathrm{d}\mathbf{r}' \int_V \mathrm{d}\mathbf{R}\, \rho(\mathbf{r}')\, \rho(\mathbf{r}' + \mathbf{R})\, E^{(i)}_m E^{(i)}_n \times \exp[i(\mathbf{k}^{(i)} - \mathbf{k}^{(s)}) \cdot \mathbf{R}]\Bigg\}. \quad (24)$$

[6] In equation (19), we switch to the usual complex notation without warning. In subsequent formulae, the left-hand side is to be interpreted as the real part of the right-hand side.

[7] The analogy is not perfect. In particular, translational invariance is preserved in equation (22). If we had a very small 'Gaussian' crystal, the electron density would have been zero outside it and the exponential factor in (23) would have been $\exp[-(|\mathbf{r}'|^2 + |\mathbf{r}' + \mathbf{R}|^2)/d^2]$.

4. The coherence of incident X-rays

The purpose of this section is to show that the proposal of this paper is a realistic one. In other words, it is fairly reasonable and simple to produce the required short mutual coherence of incident X-rays using conventional X-ray sources. I will discuss two methods to do so in some detail. First, the anode can be made relatively large and the crystal placed relatively close to it. Second, an image of the anode can be formed on the crystal using optical elements. There are undoubtedly other ways to achieve short coherence length in all three dimensions; for instance, by using synchrotron radiation. It is also clear that more detailed calculations are needed for planning real experiments.

A cautionary note: this section reviews some well known concepts in optics; therefore both trustful readers and experts can safely skip over it.

4.1. The propagation of mutual coherence

We will review now the propagation of the mutual coherence in preparation for describing some ways to minimize it at the crystal.

In §3, equation (17), we gave a general definition of the mutual coherence of the electromagnetic field. For convenience it is repeated below:

$$\Gamma^{(i)}_{m,n}(\mathbf{r}'; \mathbf{r}''; T) = \langle E^{(i)}_m(\mathbf{r}', t)\, E^{(i)}_n(\mathbf{r}'', t+T)\rangle.$$

In this section, we make further restrictive assumptions that are amply satisfied in practice. We first demand, as in §3, that the X-ray emission be a stationary random process, at least in the wide sense. Second, we assume that the X-ray spectrum incident on the crystal is narrow, $\Delta\lambda_0 \ll \lambda_0$. Third, we use a paraxial approximation that is satisfied if the opening angle of the incident X-ray beam is significantly smaller than 1 rad. As X-ray optics cannot be made today with numerical apertures much larger than 1/100, this third assumption is also well satisfied in practice. When the latter two assumptions are satisfied, we can define a center frequency ω_0 (or a center wavelength λ_0) and a dominant wavevector $\mathbf{k}^{(i)}$ that corresponds to the incident-beam direction. We will set the z direction of our coordinate system along the incident beam. In the following, we also suppress the Cartesian components m and n, assuming, in effect, scalar wave propagation.

When the (scalar) electromagnetic field, $E(\mathbf{r}', t)$, is known on a surface Σ_1 that is more or less perpendicular to the direction of propagation of the beam (the z axis), it can be calculated on another similar surface downstream from it by solving the wave equation (see Fig. 2). If the wave on the first surface is paraxial, we can either solve the paraxial wave equation in the volume between the two surfaces or, equivalently, evaluate the Huygens–Fresnel integral [for a detailed discussion the reader is referred to books by Born & Wolf (1980, ch. X) or Goodman (1985, chs. 5 and 7)]. In a similar manner, the mutual coherence, $\Gamma(\mathbf{r}; \mathbf{r}'; T)$, can also be 'propagated' from one plane to another, by evaluating the propagation of the two electric fields on the right-hand side of (17) directly. Quoting now from our sources, the mutual coherence satisfies a pair of wave equations [Goodman, 1985, equations (5.4-17, 18)]:

$$\left\{\nabla_1^2 - \frac{1}{c^2}\frac{\partial^2}{\partial\tau^2}\right\}\Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau) = 0,$$
$$\left\{\nabla_2^2 - \frac{1}{c^2}\frac{\partial^2}{\partial\tau^2}\right\}\Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau) = 0, \tag{25}$$

where the operators, ∇_1^2 and ∇_2^2 are Laplacians that operate on the two separate field points, $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively.

The alternative expression that uses Huygens's integral is equation (5.4-4) of Goodman (1985), with obliquity factors set to unity:

$$\Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau) = \int\!\!\int_{\Sigma_1}\int\!\!\int_{\Sigma_1} \frac{1}{\lambda_0 r_1}\frac{1}{\lambda_0 r_2}\Gamma\left(\boldsymbol{\rho}_1; \boldsymbol{\rho}_2; \tau + \frac{r_2 - r_1}{c}\right) \mathrm{d}S_1\, \mathrm{d}S_2, \tag{26}$$

where $\boldsymbol{\rho}_i = (\xi_i, \eta_i, \zeta_i)$, for $i = 1, 2$, denote the two 'source' points on the incident surface, and $\mathbf{r}_i = (x_i, y_i, z_i)$ are two 'field' points in the vicinity of the second surface in Fig. 2. The distances r_i are between the corresponding field and source points:

$$r_i = |\mathbf{r}_i - \boldsymbol{\rho}_i| = [(\xi_i - x_i)^2 + (\eta_i - y_i)^2 + (\zeta_i - z_i)^2]^{1/2}. \tag{27}$$

Both integrations in (26) are two-dimensional over the surface Σ_1 of the incident beam. In our notation, $\mathrm{d}S_1 = \mathrm{d}\xi_1\, \mathrm{d}\eta_1$ and $\mathrm{d}S_2 = \mathrm{d}\xi_2\, \mathrm{d}\eta_2$. Note that only the differences between the paths of the two waves, $r_1 - r_2$, enter into the propagation of the mutual coherence. Any time delay that is common to both paths cancels out.

4.2. Proximity of the anode to the crystal

Ordinary X-ray sources consist of independent atoms radiating incoherently. As the radiation propagates some distance away, it becomes coherent. One way to minimize the coherence of the X-ray beam at the crystal is to put the crystal close enough to the radiator that coherence cannot develop. We will show below, by detailed calculation, that this is a realistic proposal.

Although the sizes of the radiators in X-rays can be comparable with their wavelength, a δ-function correlation can be used in the formula. At the source the mutual coherence is then

$$\Gamma(\boldsymbol{\rho}_1; \boldsymbol{\rho}_2; \tau) = I(\boldsymbol{\rho}_1)\delta(|\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2|)\exp\left(-i2\pi\frac{c\tau}{\lambda_0}\right)\exp(-\tau\Delta\nu), \quad \tau \geq 0. \tag{28}$$

The δ function that appears, $\delta(|\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2|)$, is really two-dimensional over the radiating surface. We also assumed a narrow-band Lorentzian spectrum of width $\Delta\nu = c\Delta\lambda_0/\lambda_0^2$. In order to simplify the formulae, let us assume that the radiator is a small disc of diameter $2\rho_0$, located at the plane $\zeta = 0$, centered at $\xi = \eta = 0$, and that we are interested in the

mutual coherence of the radiation on a plane far away from it in the sense that $z \gg \lambda_0$ and $z \gg 2\rho_0$. We will also calculate the mutual coherence close to the axis, $x_1, x_2, y_1, y_2 \ll z$. We can then use the approximation, for $i = 1, 2$,

$$\begin{aligned} r_i &= [(\xi_i - x_i)^2 + (\eta_i - y_i)^2 + (\zeta_i - z_i)^2]^{1/2} \\ &\simeq z_i + [(\xi_i - x_i)^2 + (\eta_i - y_i)^2]/2z_i \end{aligned} \tag{29}$$

and, further, $z_1 \simeq z_2 \simeq z$. We can now substitute (28) into (26) and carry out one of the surface integrations. We obtain

$$\begin{aligned} \Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau) = {} & \frac{1}{(\lambda_0 z)^2} \exp(-i\psi) \exp\left[-i\frac{2\pi c}{\lambda_0}\left(\tau + \frac{z_2 - z_1}{c}\right)\right] \\ & \times \exp\left[-\frac{c\Delta\lambda_0}{\lambda_0^2}\left(\tau + \frac{z_2 - z_1}{c} + \frac{\psi\lambda_0}{2\pi c}\right)\right] \\ & \times \int\limits_{\Sigma_1}\int I(\xi, \eta) \exp\left[i\frac{2\pi}{\lambda_0 z}(\Delta x\xi + \Delta y\eta)\right] \mathrm{d}\xi\, \mathrm{d}\eta, \end{aligned} \tag{30}$$

where $\Delta x = x_2 - x_1$, $\Delta y = y_2 - y_1$ and

$$\psi = (\pi/\lambda_0 z)[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)]. \tag{31}$$

Equation (30) is called the Van Cittert–Zernike formula for a narrow-band paraxial beam.

In order to gain some understanding of (30), we shall evaluate the mutual coherence in three simple cases. The first and second are the transverse coherence of a small incoherent source. They are essentially the Fourier transforms of their intensity distribution.

First, we assume a small disc of radius ρ_0 that radiates a uniform intensity I_0. (Note that the intensity has units of E^2, the square of the electric field.) The result of the integration is

$$\begin{aligned} & \Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau) \\ & = \frac{I_0\pi\rho_0^2}{(\lambda_0 z)^2} \exp(-i\psi) \exp\left(-i2\pi\frac{c\tau}{\lambda_0}\right) \exp(-\tau\Delta\nu) \\ & \quad \times \mathbf{J}_1\left[\frac{2\pi\rho_0}{\lambda_0 z}(\Delta x^2 + \Delta y^2)^{1/2}\right] \Big/ \left[\frac{2\pi\rho_0}{\lambda_0 z}(\Delta x^2 + \Delta y^2)^{1/2}\right], \end{aligned} \tag{32}$$

where $\mathbf{J}_1[\cdot]$ is the Bessel function of the first kind of order one. We neglected the term $\psi\lambda_0/2\pi c$ in the real exponential.

Second, if we assume a Gaussian intensity distribution of the radiator,

$$I(\xi, \eta) = (I_0/2) \exp[-(\xi^2 + \eta^2)/2\rho_0^2], \tag{33}$$

the result becomes even simpler:

$$\begin{aligned} \Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau) = {} & \frac{I_0\pi\rho_0^2}{(\lambda_0 z)^2} \exp(-i\psi) \exp\left(-i2\pi\frac{c\tau}{\lambda_0}\right) \exp(-\tau\Delta\nu) \\ & \times \exp\left[-\frac{(2\pi\rho_0)^2}{2(\lambda_0 z)^2}(\Delta x^2 + \Delta y^2)\right]. \end{aligned} \tag{34}$$

In both (32) and (34), the approximate size of the transverse coherence length, d, at a distance z from the anode is $d \simeq z\lambda_0/\rho_0$, where λ_0 is the X-ray wavelength and ρ_0 is the transverse size of the X-ray emission on the anode, as seen through the collimator. A simple way of understanding this relation is to notice that, if we had a telescope of aperture d, the size of the minimum resolved spot at distance z would be given by $z\lambda_0/d$. If the actual size of the emitter is smaller than that, its different parts (that emit incoherently) cannot be resolved. Therefore, within a patch of size d, the radiator looks like a point source and its radiation is coherent. In the opposite case, if $\rho_0 \geq z\lambda_0/d$ (or, equivalently, if $d \geq z\lambda_0/\rho_0$) the beam is incoherent.

Returning now to equation (30), we calculate the longitudinal coherence of the beam, by setting $x_1 = x_2$ and $y_1 = y_2$, but $z_1 \neq z_2$ in (30) and (31):

$$\begin{aligned} \Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau) = {} & \frac{I_0\pi\rho_0^2}{(\lambda_0 z)^2} \exp(-i\psi) \exp\left[-i2\pi\frac{c}{\lambda_0}\left(\tau + \frac{z_2 - z_1}{c}\right)\right] \\ & \times \exp\left[-\frac{c\Delta\lambda_0}{\lambda_0^2}\left(-\tau + \frac{z_2 - z_1}{c}\right)\right]. \end{aligned} \tag{35}$$

Note that the longitudinal coherence of the beam depends only on its spectral width. It approximately equals $\lambda_0^2/\Delta\lambda_0$, the mean X-ray wavelength divided by its fractional spectral width.

4.3. Imaging the anode onto the crystal

The second way of making the incident X-rays relatively incoherent is by inserting optical elements that image the source onto the crystal. We use two important simplifications in our treatment. First, we model the optical system by a thin lens. Second, we assume an 'all Gaussian' optical system: a Gaussian source, as in (33), and a Gaussian aperture of the lens. The latter means that the X-ray amplitude incident on the lens is apodized (attenuated) by a Gaussian of width comparable with the lens aperture. This way we can utilize the exact formulae developed for Gaussian beams around their beam waist and the heuristic knowledge developed about them by laser physicists (Siegman, 1986). A detailed derivation is given in Appendix *A* and the interested reader is directed to it. Although X-ray imaging systems are anything but thin lenses, our simplifications lead to important conclusions that are valid to a good approximation for practical systems.

The mutual coherence in the neighborhood of the image of a source can be written as an extension of (26), by allowing the wave to propagate through an optical system. The result, quoted below, is given as equation (45) in Appendix *A*.

$$\begin{aligned} \Gamma(\mathbf{u}_1; \mathbf{u}_2; \tau) = {} & \int\limits_{\Sigma_1}\int\limits_{\Sigma_1}\int\int \mathbf{K}(\mathbf{u}_1; \boldsymbol{\rho}_1; -p_1/c)\, \mathbf{K}^*(\mathbf{u}_2; \boldsymbol{\rho}_2; -p_2/c) \\ & \times \Gamma(\boldsymbol{\rho}_1; \boldsymbol{\rho}_2; \tau)\, \mathrm{d}S_1\, \mathrm{d}S_2. \end{aligned}$$

The notation is as follows. The coordinates in the source region are denoted by $\boldsymbol{\rho}_i = (\xi_i, \eta_i, \zeta_i)$. The two 'source' points needed for the propagation of the mutual coherence correspond to $i = 1, 2$. The coordinates in the region of the image are similarly denoted by $\mathbf{u}_i = (u_i, v_i, w_i)$. The total propagation distances between the source points $\boldsymbol{\rho}_i$ and their corresponding image points $\mathbf{u}_i$ are denoted by p_i. The complex

quantity $\mathbf{K}(\mathbf{u};\boldsymbol{\rho};-p/c)$ is called the amplitude-spread function of the optical system. It is the propagation of a wave originating at a point $\boldsymbol{\rho}$ of the source to the point $\mathbf{u}$ in the image region. The principle of least time, also called Fermat's principle (Born & Wolf, 1980, p. 128), teaches us that the time delay, p/c, is well defined in the region of the geometric image of the lens. In fact, the time delay can be calculated from straight-line propagation.

For a Gaussian transmission of the lens, the amplitude-spread function can be well approximated by

$$\begin{aligned}\mathbf{K}(\mathbf{u};\boldsymbol{\rho};-p/c) = P_0\frac{\pi^{1/2}a}{\lambda_0 z_o}\exp\left[-\frac{c\Delta\lambda_0}{\lambda_0^2}\left(-\frac{p}{c}+\frac{z}{c}\right)\right]\\ \times\frac{W_0}{\tilde{q}_0 W(z)}\exp\left\{-i\left[\frac{2\pi z}{\lambda_0}-\varphi(z)\right]\right\}\\ \times\exp\left[-\frac{x^2+y^2}{W(z)^2}-i\frac{2\pi}{\lambda_0}\frac{x^2+y^2}{2R(z)}\right].\end{aligned}\quad(36)$$

The symbols in (36) are defined and discussed in Appendix *A*. For completeness, we repeat some of those definitions. The coordinates x, y and z refer to the geometrical image point and the z axis is chosen along the beam propagation direction. The beam parameters are given in (51) by

$$(\mathrm{NA}) = \frac{2^{1/2}a}{z_i};\quad W_0 = \frac{\lambda_0}{\pi(\mathrm{NA})};\quad Z_R = \frac{\lambda_0}{\pi(\mathrm{NA})^2},$$

where (NA) is the numerical aperture of the lens, W_0 is the minimum beam diameter (its waist) and Z_R, called the Rayleigh range or the depth of focus, is the length of the waist region of the Gaussian beam. In terms of these parameters, as shown in (52),

$$W(z)^2 = W_0^2\left[\left(\frac{z}{Z_R}\right)^2+1\right];\ R(z) = z+\frac{Z_R^2}{z};\ \tan[\varphi(z)] = \frac{z}{Z_R}.$$

Finally, for our restrictive assumptions of an 'all Gaussian' optical system, the mutual coherence of the beam can be evaluated explicitly in the vicinity of the focal region. To a first approximation, it is

$$\begin{aligned}&\Gamma(\mathbf{r}_1;\mathbf{r}_2;\tau)\\ &= C\frac{W_0^2}{W(z_1)\,W(z_2)}\\ &\quad\times\exp\left\{-i2\pi\frac{c}{\lambda_0}\left(\tau+\frac{z_2-z_1}{c}\right)+i[\varphi(z_2)-\varphi(z_1)]\right\}\\ &\quad\times\exp\left[-\frac{c\Delta\lambda_0}{\lambda_0^2}\left(-\tau+\frac{z_2-z_1}{c}\right)\right]\\ &\quad\times\exp\left[-\frac{x_1^2+y_1^2}{W(z_1)^2}-\frac{x_2^2+y_2^2}{W(z_2)^2}\right]\\ &\quad\times\exp\left\{-i\frac{2\pi}{\lambda_0}\left[\frac{x_1^2+y_1^2}{2R(z_1)}-\frac{x_2^2+y_2^2}{2R(z_2)}\right]\right\},\end{aligned}\quad(37)$$

where all the factors that are independent of the relative positions of $\mathbf{u}_1$ and $\mathbf{u}_2$ were lumped into the single constant C.

In plain words, the coherence volume of the image of the X-ray emitter on the crystal is essentially given by the dimensions of the image of a single point of the emitter.

The books of Born & Wolf (1980, ch. X) and Goodman (1985, chs. 5 and 7) discuss the influence of aberrations of the optical system on the coherence volume at the image. If the coherence area of the radiation incident on the optical system is much smaller than the input pupil (aperture) of the lens, the lens is illuminated incoherently. The condition for incoherent illumination is that the lens is large enough and close enough to the source [see equation (34)]. In this case, the output aperture of the lens acts as another incoherent source and the coherence volume around the geometrical image can again be calculated by equation (34). An important corollary is that the coherence volume is quite insensitive to aberrations of the optical system. Note that the illuminated area of the lens is generally much larger than the X-ray source itself, so the lens acts to make the coherence volume small.

This section can be summarized by three important 'rules of thumb'. First, the illuminated area of the crystal is approximately the size of the emitting anode multiplied by the magnification of the optical elements (a result from geometrical optics). Second, the mutual coherence of the beam at the crystal face is similar to that obtained from an incoherent source covering the exit pupil (aperture) of the focusing lens. Thus, its area is usually much smaller than the illuminated area. In particular, if the optical system has a numerical aperture (NA), the transverse coherence length is approximately $\lambda_0/(\mathrm{NA})$. The longitudinal coherence length is the smaller of $\lambda_0/(\mathrm{NA})^2$ and $\lambda_0^2/\Delta\lambda_0$, the latter limit given by the spectral width of the incident X-rays. Third, the coherence volume is relatively insensitive to aberrations in the X-ray optics.

5. 'Solving the phase problem'

In the *Introduction*, I promised to present a new physical method for a partial solution of the crystallographic phase problem. In order to justify such a weighty statement of intent, I will describe here three different approaches to the phase problem and discuss them in some detail.

First, it has been known for more than 90 years that the amplitudes of Bragg reflections from crystals can be measured but their phases cannot. Therefore, by crystallographic tradition, any prior or physical information about the crystal is viewed as phase information.

A second approach considers the solution of the crystal structure to be an 'inverse problem': the diffraction pattern produced by the crystal is observed and our task is to find the source of the diffracted waves. From this point of view, the phase ambiguity corresponds to different scatterers that produce the same diffraction intensity, wherever the latter is measured.

A third way, which generalizes the second approach, is through the analogy of X-ray crystallography and holography (Tollin *et al.*, 1966; Szöke, 1993). From this point of view, the recovery of the crystal structure is analogous to the holographic inverse problem: we try to find the unknown object whose hologram was observed. The ambiguity of the recov-

ered crystal structure corresponds then to the holographic dual image.

I will finish the section by discussing algorithms to find the electron density. In particular, I will describe how our 'holographic' method in crystallography, *EDEN*, was adapted to recover the electron density from an (oversampled) continuous diffraction pattern, not dissimilar to the one proposed in this paper.

5.1. 'The missing phases'

To repeat, it was well known that the missing phases were a source of difficulty in obtaining accurate crystal structures. Soon after the announcement of the sampling theorem by Shannon (1949), it was shown by Sayre (1952) that the Bragg conditions are equivalent to a critical sampling of the diffraction pattern. This established once and for all that it is lack of information that prevented the unique solution of crystal structures from measured diffraction data and not the lack of skill of the crystallographers.

The sampling theorem of Shannon ensures that a function of finite bandwidth is completely determined by its sampled values if they are sampled closely enough [for a discussion see work by Bricogne (1992, §1.3.1.6.1)]. Equivalence of the Bragg conditions to critical sampling means that, if the phases of crystal reflections are known, the electron density can be reconstructed everywhere; but any missing information makes the solution of a crystal structure ambiguous. The subject is further clarified by Daubechies (1992, ch. 2) who teaches us that undersampling causes ambiguity, critical sampling produces a representation of the electron density that converges very slowly and oversampling allows a rapidly convergent representation. Therefore, more information always helps. I quote Bricogne's remarks on this matter (Bricogne, 1992, §1.3.3.1.1.7): 'Thus the loss of phase is intimately related to the impossibility of intensity interpolation, implying in return that any indication of intensity values attached to non-integral points of the reciprocal lattice is a potential source of phase information'.

While there is general agreement in the literature on the equivalence of Bragg's conditions and critical sampling, there is no consensus on the question of how much information is sufficient to compensate for the missing phases of the reflections. This clearly depends on the kind of obtainable information. For example, it is generally accepted that two independent derivatives in multiple isomorphous replacement, or a single derivative and one-wavelength anomalous dispersion, produce enough information to solve the crystal structure. With non-crystallographic symmetry, which is equivalent to obtaining reflection intensities at non-integral values of the reciprocal lattice, it is less clear whether a twofold or threefold symmetry is enough to solve the structure. The use of spatial information seems to be very effective. With the holographic method, *EDEN*, we succeeded in solving the crystal structure from a simulated (noiseless) diffraction pattern and from the position of empty regions in 60% of the unit cell. Béran & Szöke (1995), using more sophisticated methods, have recovered a similar crystal from the knowledge of a little more than 50% of the contents of the unit cell, most of which was in the empty region. When high-resolution data are available, direct methods can recover the whole crystal *ab initio*.

5.2. Diffraction, the inverse problem and its null space

The solution of a crystal structure is an 'inverse problem': the diffraction pattern produced by the crystal is observed and our task is to find the source of the diffracted waves. Such inverse problems have been studied by mathematicians and they are known to be very difficult, unstable and ill posed (Bertero, 1989; Luenberger, 1984; Sabatier, 1987; Natterer, 1986). In plain language, this means that there are usually many crystal structures that produce almost the same diffraction patterns and that very small measurement errors correspond to very large differences in the crystal structure. It cannot be repeated enough times that these are properties of the problem itself and that no mathematical magic can undo them.

The Bragg condition prevents the measurement of the diffraction pattern everywhere. This leads to inadequate sampling of the diffraction pattern and additional uncertainty in the recovery of the scatterer, as discussed above.

Nevertheless, even if the diffraction pattern were measured everywhere, our difficulties would not yet be over. There are still several 'trivial' ambiguities in the solution of the inverse problem (Miao *et al.*, 1998). The first one is that the origin of the unit cell can be placed at an arbitrary position: two electron densities, $\rho(\mathbf{r})$ and $\rho(\mathbf{r}+\mathbf{r}_0)$, with any constant $\mathbf{r}_0$ give the same diffraction pattern. The second one is the enantiomorph: two crystal structures that have the densities $\rho(\mathbf{r})$ and $\rho(-\mathbf{r})$ are centric reflections of each other; if the densities are real numbers, they also have the same diffraction pattern. Note that anomalous scattering is capable of resolving the ambiguity. The third one is the Babinet opposite: two electron densities, $\rho(\mathbf{r})$ and $C-\rho(\mathbf{r})$ for any real constant C have the same diffraction pattern except the $(h,k,l)=(0,0,0)$ term that cannot be measured. Theoretically, it is important to realize that two crystal structure solutions that relate to each other by these symmetries are equally valid.[8]

5.3. The holographic inverse problem

When Gábor invented holography (Gábor, 1948, 1949), he paved the way to obtain three-dimensional images at atomic resolution. Stated in general terms, his insight was that, when a known wavefront is made to interfere with an unknown wavefront, the phase of the latter becomes measurable (with twofold uncertainty). He proposed a two-step method to recover three-dimensional objects using this phase information. In the first step, the recording of the hologram, the

[8] In practice, we have observed many times that some mixtures of enantiomorphic structures were obtained by our crystallographic recovery program, *EDEN*. This is a well known problem in image-processing phase retrieval (Stark, 1987). In an 'easy' case, where difficulties of this sort do not arise, *EDEN* succeeded in the model-free recovery of the electron density of a silicon crystal *ab initio*.

unknown object and a reference object are illuminated by a coherent source of waves. In its simplest version, the reference object is a small (point-like) scatterer. The waves scattered from the reference object and those scattered from the unknown object form an interference pattern that is recorded on a screen at some distance away. Gábor called this recording a hologram. The name implies that there is three-dimensional 'total' information contained in the recorded interference pattern. The second step is the reconstruction of the scattering object from this recording. For this step, he proposed to illuminate the hologram with a replica of the reference wave and showed that part of the wave transmitted by the hologram is the same as the wave originally scattered by the unknown object. Three-dimensional reconstruction by holography is so widespread and successful that most of us do not even stop to contemplate the holograms we encounter daily on our credit cards, drivers' licences, passports and even paper currency.

Gábor's method of reconstruction corresponds to the way we see a luminous (or illuminated) object: we extrapolate the waves coming from the object backwards and we place the object at the position where those waves are strongest. Barton (1991) developed the method for computers and it was used very successfully for the reconstruction of three-dimensional images of atomic resolution both from X-ray and from electron holograms. Recent achievements are described in a paper by Tegze *et al.* (2000) and in the references they cite.

The close connection between holography and crystallography has been recognized many times. In his original paper, Gábor credited Bragg's optical ideas for sparking the invention of holography. Some years later, the eminent holographer George Stroke and the eminent crystallographers Peter Main and Michael Rossmann (Tollin *et al.*, 1966) discussed the subject very clearly. I took it up some years ago (Szöke, 1986, 1993) and, naturally, I will discuss it from my point of view.

When the analogy of holography and crystal diffraction is recognized, the 'solution' of crystal structures is analogous to the holographic inverse problem. The easiest way to see that is to imagine that part of the crystal structure is known and that this known part of the structure serves as the reference wave of the hologram. This leads to the generalization of the diffraction inverse problem and it obviously reduces to it when no part of the scatterer is known.

There are two origins to the ambiguity in holographic reconstruction: one is called the holographic dual image, the other one is the lack of adequate sampling (Szöke, 1993). A hologram is really an interferogram of a sort. When two waves interfere, the resulting intensity depends on the cosine of their phase difference. As the cosine of an angle is the same as the cosine of the negative angle, interference patterns have an 'inherent' twofold ambiguity. The holographic dual image is a (very complicated) elaboration of this fact. Coming now to the question of adequate sampling, the possibility and necessity of oversampling was recognized very early in holography (Gábor, 1949; Leith & Upatnieks, 1962). Therefore there has been very little explicit discussion of it in recent literature. Naturally, in discussing crystallography from a holographic perspective, the subject has been treated quite extensively. For example, in a recent paper (Szöke, 1997) I gave a very detailed discussion of the properties of the holographic dual image; over- and undersampling; and the stability of the holographic recovery. I refer the interested reader to that paper for details.

5.4. Recovery algorithms

I will mention briefly three different but closely related algorithms to 'solve' an oversampled diffraction pattern. The first one is based on Gábor's algorithm (Barton, 1991), as mentioned above. It is used very successfully in X-ray holography and photoelectron holography. It is closely related to the back-projection method in tomography (Natterer, 1986). I would like to point out that the coverage of the reciprocal space is very sparse even in multiple-energy X-ray holography; yet, with oversampling, Barton's algorithm recovers the scatterers to high resolution and fidelity. The second one follows the tradition of 'phase retrieval' algorithms in astrophysics and image processing. It essentially iterates between real and reciprocal space, enforcing the appropriate constraints at each step. Recent work by Miao & Sayre (2000) gives a clear review of the subject and promises to be able to recover the molecule from an oversampled diffraction pattern.

The third algorithm is a crystallographic recovery algorithm that is based on the connection between holography and crystallography, which we have been developing for some years now (Szöke, 1993; Szöke *et al.*, 1997; Szöke, 1998). The program *EDEN* focuses on recovering the electron density of the crystal using all available information. Of course, there is a unique relation between the electron density and the complex structure factors, but we have found it much more illuminating and useful to treat real-space information in real space and reciprocal-space information in reciprocal space.

In a recent paper (Szöke, 1999), I discussed X-ray diffraction from a crystal that contains two similar molecules. I assumed that the molecules occupy random positions: any position in the crystal has the same probability of being occupied by one or the other kind of molecule. I have shown that the diffraction pattern of such a crystal has two components. The usual Bragg pattern is that of the weighted average of the two molecules. The second component is a continuous (diffuse) diffraction pattern; it is that of the 'difference' molecule. The overall intensity of the continuous diffraction pattern is, of course, maximum when half the molecules are of one kind and half of them of the other kind. We simulated this continuous diffraction pattern using two states of the photoactive yellow protein as deposited in the Protein Data Bank at that time. We placed the difference molecule into an artificial unit cell that is twice the size of the real unit cell. This allows the diffraction to be sampled at double the critical frequency in all three spatial dimensions. We also constrained the molecule to be in only one of the eight original unit cells (by demanding that the electron density be zero in the other 7/8 of the artificially doubled unit cell). Our program *EDEN* recovered the electron density of the difference molecule essentially without error, when some partial information was supplied as a starting point for recovery. The paper gives a

rather detailed discussion of the uniqueness, stability and convergence of the recovery in terms similar to those outlined above.

Recently, David Sayre and his collaborators (Miao *et al.*, 1999) measured the diffraction pattern of a small (micrometre scale) planar object. They used oversampling and an iterative 'phase retrieval' algorithm (mentioned above) to recover the object. Their work brought new focused attention to the possibility of 'solving the phase problem' by oversampling the diffraction pattern. They also coined the phrase 'oversampling method' (Bates, 1982) that aptly describes much of what is written in the present paper.

6. Summary and discussion

I summarize here our main results. In §3, we derived expressions for the scattered intensity of a partially coherent X-ray beam incident on a crystal. Our most general result was that, in any particular direction, the scattered intensity is proportional to the Fourier transform of the electron-density autocorrelation function of the crystal multiplied by the mutual-coherence function of the incident beam [equation (18)]. This was the point of departure for the rest of the paper. For the usual coherent incident X-ray beam, in the kinematic approximation, the scattered intensity is proportional to the Fourier transform of the electron-density autocorrelation function, without the second factor. We surmised that if the coherence length of the incident X-ray beam could be kept almost as small as a unit cell, the usual diffraction peaks would smear out and, therefore, we could obtain information on the electron density in directions that do not necessarily satisfy the Bragg conditions.[9]

Although the derivations in this paper were rigorous, I also tried to give intuitive explanations of my ideas at each stage. In case I have not succeeded, I would like to address here several nagging questions. First, are there other ways to derive the same results, using more familiar concepts? Second, we know that, except for spectral regions of anomalous dispersion, changing the wavelength of the incident X-rays does not give independent information because measurable intensities are always along reciprocal-lattice vectors of the crystal. Can we justify in simple terms that an incident beam of limited coherence gives information on Fourier components of the unit cell in directions where the electron density is not periodic in the crystal? Finally, why do the Bragg reflections lose so much of their intensity under these conditions?

The electric field amplitude of the scattered wave is a linear function of the incident field amplitude and, therefore, of the field emitted by the X-ray source. In the main body of the paper we calculated the observed diffraction intensity; a quantity that is quadratic in the electric field. The second-order non-linearity of the detector introduced the Patterson function of the crystal as well as the mutual coherence of the incident field. Here I will show, at least in outline, that the scattered field itself can be calculated in two different ways. The first way is to calculate the incident field at the position of the crystal as a sum of the fields emitted by the (independently) radiating X-ray emitters; then, starting from this partially coherent incident field, calculate the scattered field. The second way is to start from each radiating atom, calculate the field scattered by the crystal, then add up the scattered-field amplitudes for the independently radiating atoms. We should get the same result as in the main body of the paper.

Below, I will calculate the scattered field for an arbitrary vector of reciprocal space and show that its amplitude is finite, smoothly varying with direction and that its square is proportional to the number of unit cells in the crystal. I will also show that the coherence criteria I proposed are necessary and sufficient for those properties.

The first type of calculation paraphrases the derivation in the main body of the paper in terms of electric field amplitudes instead of their mutual coherence. Let us assume that the coherence volume of the incident electric field coincides with the unit cell of the crystal. Let us concentrate on a particular unit cell of the crystal. The incident electric field is the sum of all the fields radiated by all atomic radiators and propagated to the particular unit cell. It is a sum of many vectors with random amplitudes and phases; therefore its complex amplitude is a Gaussian random vector. The field is coherent within the unit cell, meaning that whatever the amplitude and phase of the field is at one point, it is about the same at other points of the unit cell. The scattered field can now be calculated from that particular unit cell; it is proportional to the incident field, but its phase is shifted depending on the exact position of the unit cell that does the scattering. (Of course, the scattered field is that of a non-periodic object, therefore it is continuous.) We can repeat the calculation for other unit cells. As the incident fields have Gaussian random amplitudes and uniform random phases, so does the scattered field. The result is that the sum of all the scattered fields is a Gaussian random field with an amplitude proportional to $N^{1/2}$, where N is the number of unit cells in the crystal. The time-averaged intensity is, therefore, N times the diffraction intensity of the single unit cell. Note that the relative phase of the diffracted field from two different unit cells is always random because the incident field has random phases. In Bragg directions, the relative phase shift of the scattered field with respect to the incident field from two different unit cells is $2n\pi$, so the electric fields add in phase if the incident fields have the same phase. This is the reason that, usually, Bragg reflections have strong intensities. Under our assumption of limited coherence, the scattered field is still a sum of Gaussian random vectors even in the directions of reciprocal-lattice vectors.

The second calculation is more subtle. Let us consider an individual emitter. The field incident on the crystal is coherent over many unit cells. Therefore, our expectation is that the radiation scattered from the crystal should be restricted to the Bragg directions. However, if we account for the randomness

[9] We enumerated the assumptions that led to equation (18) rather carefully. Out of the long list, the limiting assumptions are: the incident X-ray spectrum is narrow, $\Delta\lambda_0/\lambda_0 \ll 1$; there is no anomalous scattering; only the single-scattering (kinematic) approximation is considered; the crystal is non-magnetic and non-conducting. The other assumptions are usually very well satisfied.

of the scattered-field amplitudes from the different emitters, we get the same result as in the first calculation.

Some details of the calculation are given in Appendix *B*; I will outline the main steps here. Individual atomic emitters are modeled as point radiators with Gaussian random amplitudes and random phases which start to radiate at uniformly random times and decay in a short time, each radiator emitting at its own frequency. It is assumed that the whole emitting region is 'small' and the radiation is 'narrow band', so that there is a main incident wavevector, $\mathbf{k}^{(i)}$, and the rth individual radiator's wavevector is given by a small deviation from it, $\mathbf{k}_r^{(i)} = \mathbf{k}^{(i)} + \Delta\mathbf{k}_r^{(i)}$. The electric field at the position $\mathbf{r}'$ and time t' in the crystal is then calculated and equation (12) is used to calculate the diffracted field in the arbitrary direction, $\mathbf{k}^{(s)}$. Next, the sum over all the emitters is carried out. We notice that an 'equivalent' electron density can be defined, given by the real electron density, $\rho(\mathbf{r}')$, multiplied by phase terms that originate from the deviation of individual emitters from the average direction, as well as from the deviations of their frequency from the average. The scattered field is the diffraction pattern of the equivalent electron density, integrated over the crystal. Closer examination shows that, if the conditions outlined in the paper are satisfied, the equivalent electron density acquires a random phase in each unit cell and the results of the paper are recovered. In some sense we transferred the coherence of the field into an equivalent electron density. In Appendix *B*, I also show that the conditions are necessary and I digress on an alternative interpretation of convergent-beam monochromatic diffraction and of narrow-band Laue diffraction.

The purpose of §4 was to show that the mutual-coherence volume of the incident X-rays can be made small with very reasonable means and effort. All present X-ray sources are composed of incoherent emitters but, as X-rays propagate away from the source, they become more and more coherent. I outlined two ways to minimize the coherence of the X-ray beam incident on the diffracting crystal. First, the source of X-rays (the anode) can be made relatively large and the crystal can be placed relatively close to it. Second, an image of the anode can be formed on the crystal using optical elements.

The simpler way is the first one. We used the Van Cittert–Zernike formula for a narrow-band paraxial beam [equation (30)] to show that the transverse coherence of the incident X-ray beam is approximately $z\lambda_0/r$, where λ_0 is the X-ray wavelength and r is the transverse size of the X-ray emitter, at a distance z away, as seen through the collimator. For instance, for an emitting anode diameter of $r = 0.1$ cm, a mean X-ray wavelength of $\lambda_0 = 1.5$ Å and a distance of $z = 5$ cm to the crystal, we obtain a transverse coherence length $d \simeq 75$ Å, which can be comparable to the size of the unit cell of a macromolecular crystal. Note that the transverse coherence length decreases if we increase the size of the emitter, bring the crystal closer to the anode and use a shorter X-ray wavelength. The longitudinal coherence, given by equation (35), is approximately $\lambda_0^2/\Delta\lambda_0$, the X-ray wavelength divided by its fractional spectral width. In order to obtain a longitudinal coherence length of ~75 Å, one has to have a fractional bandwidth of ~0.02. This excludes the use of characteristic radiation, as it usually has an intrinsic bandwidth of $\leq 10^{-4}$.

The second method is to form an image of the anode at the crystal, using optical elements. We established three important 'rules of thumb' (Born & Wolf, 1980, ch. X; Goodman, 1985, chs. 5 and 7). First, the illuminated area of the crystal is approximately the size of the emitting anode multiplied by the magnification of the optical elements, as a result of geometrical optics. Second, the mutual coherence of the beam at the crystal face is similar to that obtained from an incoherent source covering the exit pupil of the focusing lens. Thus its area is usually much smaller than the illuminated area. In particular, if the optical system has a numerical aperture (NA), the transverse coherence length is approximately $\lambda_0/(\mathrm{NA})$. The longitudinal coherence length is the smaller of $\lambda_0/(\mathrm{NA})^2$ and $\lambda_0^2/\Delta\lambda_0$, the latter limit given by the spectral width of the incident X-rays. Third, the coherence volume is relatively insensitive to aberrations in the X-ray optics.

An experiment that suggests itself is to move a suitable crystal along the axis of the imaging optics. When the crystal is far from the focus, we should be able to record the usual Bragg diffraction pattern. As the crystal moves into the focal region, the diffraction pattern should broaden and, if the conditions are right, it should disappear. One obvious advantage of such an experiment is that the diffuse background, resulting from thermal motion and crystal imperfections, can be monitored and subtracted properly. It is clear from the preceding considerations that more careful and detailed theoretical development is needed for planning real experiments. It is of particular importance to consider synchrotron sources, where the situation is more complicated.

In §5, I discussed three appproaches to the 'phase problem' of crystallography. Traditionally, crystallographers think in terms of reciprocal space. I described an alternative point of view that is a relative of holography. I found it fruitful for finding new methods to obtain information on electron densities as well as for the consistent and efficient use of available information. In particular, our program *EDEN* is capable of recovering the electron density from an oversampled diffraction pattern, given a suitable starting point.

Some remarks are in order about previous similar work. The closest one is that of Dušek (1994, 1995) who calculated the diffraction from a crystal taking into account the mutual coherence of the incident X-rays. His formulae are, of course, identical to ours. Unfortunately, he did not consider the case where the unit-cell dimensions of the crystal are so much larger than the X-ray wavelength that the coherence length of the incident radiation can be made comparable with them. His work concentrated on small shifts in the diffraction maxima that occur when the transverse coherence length of the beam is not much larger than the wavelength.

In this paper, I used material only from venerable optics textbooks and papers. After all, X-rays are electromagnetic waves. I did not consider similar literature on electron microscopy. I have no doubts that analogous experiments can be carried out using light microscopes, soft X-ray microscopes

and electron microscopes (Cowley, 1995; Pendry, 1974). I would not be surprised to learn that such experiments have indeed been performed and I am ready to concede priority to anyone who did them.

APPENDIX *A*
The propagation of mutual coherence through X-ray optics

In this Appendix, we develop formulae for the propagation of the mutual coherence of X-rays from an incoherent source, through a thin lens, to the region of the image. We follow closely Goodman's (1985) treatment. The notation is that of Fig. 3(*a*). We need three coordinate systems. The coordinates in the source region are denoted by $\boldsymbol{\rho}_i = (\xi_i, \eta_i, \zeta_i)$, where $i = 1, 2$ denote the two 'source' points needed for the propagation of the mutual coherence. The two similar points in the region of the lens are $\mathbf{r}_i = (x_i, y_i, z_i)$ and the distances between $\boldsymbol{\rho}_i$ and $\mathbf{r}_i$ are r_i, as defined in equation (27). The third coordinate system is in the region of the image, where the diffracting crystal is to be placed. The two image points in it are denoted by $\mathbf{u}_i = (u_i, v_i, w_i)$. The total propagation distances between the source points, $\boldsymbol{\rho}_i$, and their corresponding image points, $\mathbf{u}_i$, are denoted by p_i. They are defined similarly to equation (27).

The mutual coherence at the input plane of the thin lens is given by (30) and (31) with $z_1 = z_2 = z = z_o$, the object distance. Note, specifically, that ψ in (31) describes the curvature of the wavefront; it is that of a spherically divergent wave centered at the source. For completeness, we repeat equations (30) and (31) with the curvature parameter denoted by ψ_o, emphasizing that it is the curvature of the wave emanating from the source:

$$\Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau)_{\rm in} = \frac{1}{(\lambda_0 z_o)^2} \exp(-i\psi_o) \exp\left(-i\frac{2\pi\tau c}{\lambda_0}\right) \times \exp\left[-\frac{c\Delta\lambda_0}{\lambda_0^2}\left(\tau + \frac{\psi_o\lambda_0}{2\pi c}\right)\right] \times \int\!\!\int_{\Sigma_1} I(\xi, \eta) \exp\left[i\frac{2\pi}{\lambda_0 z_o}(\Delta x\xi + \Delta y\eta)\right] {\rm d}\xi\, {\rm d}\eta,$$

where $\Delta x = x_2 - x_1$, $\Delta y = y_2 - y_1$ and

$$\psi_o = (\pi/\lambda_0 z_o)[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)].$$

The thin lens will be described by an amplitude transmission

$$P(x, y) = P_0(x, y) \exp[-(x^2 + y^2)/2a^2], \qquad (38)$$

where a, of the order of the radius of the lens, denotes its 'Gaussian apodizing radius'. $P_0(x, y)$, the 'aperture function' of the lens, is a slowly varying, possibly complex, function, the complex part of which describes the aberrations of the lens. The lens also introduces a time delay that depends on its focal length, f, and the radial coordinates of the (paraxial) ray:

$$\delta\tau = -(x^2 + y^2)/2cf. \qquad (39)$$

If we put the time delay into (30), we obtain

$$\Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau)_{\rm out} = \Gamma(\mathbf{r}_1; \mathbf{r}_2; \tau)_{\rm in}\, P(x_1, y_1)\, P^*(x_2, y_2) \times \exp\left\{i\frac{\pi}{\lambda_0 f}[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)]\right\} \times \exp\left[-\frac{c\Delta\lambda_0}{\lambda_0^2}\frac{(\psi_i - \psi_o)\lambda_0}{2\pi c}\right], \qquad (40)$$

where ψ_i is defined in (44) below. We can substitute (40) into (30), then use the propagation formula (26) to obtain the mutual coherence at any point in space.

The formulae can be considerably simplified in our 'all Gaussian' case. For a start, we multiply the two curvature terms in (30) and (40):

$$\exp(-i\psi_o) \exp\left\{i\frac{\pi}{\lambda_0 f}[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)]\right\} = \exp\left\{i\left(\frac{\pi}{\lambda_0 f} - \frac{\pi}{\lambda_0 z_o}\right)[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)]\right\} = \exp\left\{i\frac{\pi}{\lambda_0 z_i}[(x_2^2 + y_2^2) - (x_1^2 + y_1^2)]\right\}. \qquad (41)$$

In order to simplify the right-hand side, we used the connection between the object distance, z_o, the image distance, z_i, and the focal length, f, for the thin lens:

$$1/z_o + 1/z_i = 1/f. \qquad (42)$$

Equation (41) shows that, at the output of the lens, a spherical wavefront centered at the image point is obtained. (This is, of course, the condition for image formation.) Let us write explicitly

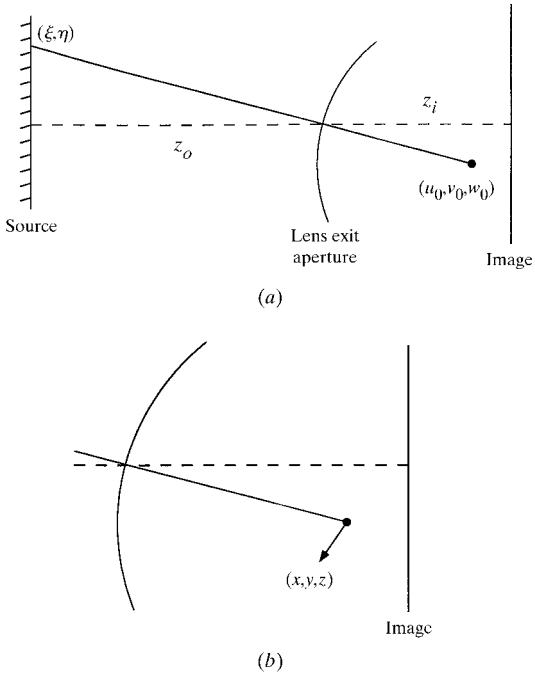

Figure 3
(*a*) Geometry for imaging an incoherent source onto a crystal. (*b*) Details in the vicinity of the image.

$$\Gamma(\mathbf{r}_1;\mathbf{r}_2;\tau)_{\rm out}=\frac{1}{(\lambda_0 z_o)^2}\,P(x_1,y_1)\,P^*(x_2,y_2)\times\exp(-i\psi_i)\exp\left(-i\frac{2\pi\tau c}{\lambda_0}\right)\times\exp\left[-\frac{c\Delta\lambda_0}{\lambda_0^2}\left(\tau+\frac{\psi_i\lambda_0}{2\pi c}\right)\right]\times\int\!\!\int_{\Sigma_1} I(\xi,\eta)\exp\left[i\frac{2\pi}{\lambda_0 z_o}(\Delta x\xi+\Delta y\eta)\right]\mathrm{d}\xi\,\mathrm{d}\eta, \tag{43}$$

where

$$\psi_i=-(\pi/\lambda_0 z_i)[(x_2^2+y_2^2)-(x_1^2+y_1^2)]. \tag{44}$$

At this point, we could use equation (7.1-46) of Goodman (1985) to obtain the mutual coherence of the X-ray beam at the crystal. The result would be similar to Goodman's 'formidable result', equation (7.1-36).

A much simpler formula can be obtained for the propagation of a Gaussian X-ray beam through an 'all Gaussian' optical system. We are guided by the general result that the mutual coherence of the radiation in the image region can be decomposed, as in equation (7.1-39) of Goodman (1985):

$$\Gamma(\mathbf{u}_1;\mathbf{u}_2;\tau)=\int\!\!\int_{\Sigma_1}\int\!\!\int_{\Sigma_1}\mathbf{K}(\mathbf{u}_1;\boldsymbol{\rho}_1;-p_1/c)\,\mathbf{K}^*(\mathbf{u}_2;\boldsymbol{\rho}_2;-p_2/c)\times\Gamma(\boldsymbol{\rho}_1;\boldsymbol{\rho}_2;\tau)\,\mathrm{d}S_1\,\mathrm{d}S_2. \tag{45}$$

The complex quantity $\mathbf{K}(\mathbf{u};\boldsymbol{\rho};-p/c)$ is called the amplitude spread function of the optical system. It describes the propagation of a wave originating at a point $\boldsymbol{\rho}$ of the source to the point $\mathbf{u}$ in the image region. We will restrict the transmission factors to be Gaussian, as in (38).

Let us choose a single point of the source, $\boldsymbol{\rho}=(\xi,\eta)$. Let us also choose one of the coordinates in the region of the lens, $\mathbf{r}_1=(x_1,y_1)$. The part of equation (43) that will contribute to the amplitude spread function, $\mathbf{K}(\mathbf{u}_1;\boldsymbol{\rho}_1;-p/c)$, is given by the factor

$$P_0(x_1,y_1)\exp\left(-\frac{x_1^2+y_1^2}{2a^2}\right)\exp\left[-i\frac{\pi}{\lambda_0 z_i}(x_1^2+y_1^2)\right]\times\exp\left[-i\frac{2\pi}{\lambda_0 z_o}(x_1\xi+y_1\eta)\right]. \tag{46}$$

The amplitude spread function can then be calculated, once the additional propagation of the wave between the lens and the focal region is taken into account.

The wavefront emerging from the lens is constant on a spherical surface that is centered on the geometric image point of the source point, (ξ,η). Referring to Fig. 3(*b*), the geometric image point is

$$(u_0,v_0,w_0)=\left(-\frac{z_i}{z_o}\xi,-\frac{z_i}{z_o}\eta,-\frac{u_0^2+v_0^2}{2z_i}\right), \tag{47}$$

as can be seen by demanding that the image point has the distance z_i from the center of the lens and expanding the resultant expression to first order. If we now draw a sphere around (u_0,v_0,w_0) with radius z_i, the coordinates of this sphere, in the lowest-order approximation, are the same as the complex factors in (46). We assume that the thin lens is aberration free.

If we assume that $P_0(x,y)$ is approximately constant, the wave emerging from the lens is a Gaussian spherical wave (Siegman, 1986, §16.3, equation 33). More accurately, its complex amplitude is that of a lowest-order Gaussian beam far from its focus. In order to simplify our notation, we describe the propagation of the beam in a coordinate system where the beam propagates along the z axis and we put the origin of the coordinate system at the geometrical image point. The propagation of such a wave under paraxial conditions can be written in closed form, either by solving the paraxial form of the wave equation, (25), or by evaluating the Huygens integral, (26). The result is self-similar propagation in the sense that the wave stays Gaussian everywhere; only its diameter and its radius of curvature change. An 'elegant' way of representing the propagation of a Gaussian beam is the following. Let us define a (properly normalized) scalar wave amplitude,

$$E(x,y,z)=\left(\frac{2}{\pi}\right)^{1/2}\frac{W_0}{\tilde{q}_0 W(z)}\exp\left\{-i\left[\frac{2\pi z}{\lambda_0}-\varphi(z)\right]\right\}\times\exp\left[-\frac{x^2+y^2}{W(z)^2}-i\frac{2\pi}{\lambda_0}\frac{x^2+y^2}{2R(z)}\right]. \tag{48}$$

The parameters $W(z)$ and $R(z)$ denote the diameter and the radius of curvature of the wave, respectively. The z-dependent phase shift, $\varphi(z)$, is defined in equation (52) below. It is important only in the focal region. We also define a complex beam radius, $\tilde{q}(z)$, by

$$1/\tilde{q}(z)=1/R(z)-i\lambda_0/\pi W(z)^2. \tag{49}$$

Then the entire propagation of the beam is expressed by the simple equation

$$\tilde{q}(z)=\tilde{q}_0+z. \tag{50}$$

The parameters appearing in equations (48), (49) and (50) can be connected with those appearing in (47) by assuming that the diameter of the lens is much larger than that of the focused beam. In fact, the beam parameters can all be calculated from the numerical aperture of the optical system, using the following relations. We first define three parameters:

$$(\mathrm{NA})=\frac{2^{1/2}a}{z_i};\quad W_0=\frac{\lambda_0}{\pi(\mathrm{NA})};\quad Z_R=\frac{\lambda_0}{\pi(\mathrm{NA})^2}. \tag{51}$$

(NA) is the numerical aperture of the lens, W_0 is the minimum beam diameter (its waist) and Z_R, called the Rayleigh range or the depth of focus, is the length of the waist region of the Gaussian beam. In terms of these parameters,

$$W(z)^2=W_0^2\left[\left(\frac{z}{Z_R}\right)^2+1\right];\quad R(z)=z+\frac{Z_R^2}{z};\quad \tan[\varphi(z)]=\frac{z}{Z_R}. \tag{52}$$

Around the beam waist, at $z=0$, the minimum beam diameter is W_0. It is clear from (47) that the beam diameter

stays within a factor of $2^{1/2}$ of its minimum as long as $|z| \leq Z_R$. The wavefront of the beam flattens out around $z = 0$, as $R(z) \rightarrow \infty$ there. At large distances from the focus, $|z| \gg Z_R$ and $R(z) \rightarrow z$, with an error of Z_R^2/z. The so-called Guoy phase shift, $\varphi(z)$, causes a lengthening of the wavelength in the focal region. The lengthening at $z = 0$ is given by $\delta\lambda_0/\lambda_0 = \lambda_0/Z_R$. The more general formulae (52) and (48) show that at $z = \pm Z_R$ the phase shift is $\pm 45^\circ$ and a full passage through the focus always causes a 180° total phase shift of the wave.

APPENDIX *B*
Calculation of the diffracted field

In this Appendix, the calculation of the diffracted field from a crystal without the use of correlation functions is outlined. The X-ray source will be modeled as a collection of independent point radiators and the scattered field for a general direction of the detector will be calculated. I will attempt to show that the diffraction pattern is the incoherent sum of those of individual molecules if and only if the coherence volume of the radiation is of the order of that of the unit cell in all three dimensions. I also add some remarks on Bragg scattering using a convergent beam and on 'narrow band' Laue diffraction.

We start from equation (12) for calculating the scattered electric field at the detector, $E^{(s)}(\mathbf{r}, t)$, and make the following simplifications. We use complex notation and consider scalar fields. We also ignore disorder and thermal motion in the crystal and use its electron density, $\rho(\mathbf{r})$, for the susceptibility. The result is

$$E^{(s)}(\mathbf{r}, t) = \frac{k_0^2}{r^{(s)}} \int_V \rho(\mathbf{r}')\, E^{(i)}(\mathbf{r}', t' + \mathbf{r}' \cdot \hat{\mathbf{s}}/c)\, \mathrm{d}\mathbf{r}', \qquad (53)$$

where the symbols are the same as in Fig. 1. The reader is reminded that $\mathbf{r}$ and t refer to the region of the detector, $\mathbf{r}'$ and t' refer to the crystal region, and $\mathbf{r}''$ and t'' refer to the region of the X-ray source. The superscripts (i) and (s) indicate the incident and scattered waves, respectively. The relations among the coordinates of the detector and the crystal are $\mathbf{r} = r^{(s)}\hat{\mathbf{s}} = r^{(s)}(\mathbf{k}^{(s)}/k^{(s)})$, where $\hat{\mathbf{s}}$ is a unit vector in the direction of the scattered wave, $\mathbf{k}^{(s)}$ is the wavevector of the scattered wave and $k^{(s)}$ is its magnitude. The average retardation of the wave at the crystal is $t' = t - r^{(s)}/c$. The distance from the crystal to the detector is approximated as $|\mathbf{r} - \mathbf{r}'| \simeq r^{(s)} - \mathbf{r}' \cdot \hat{\mathbf{s}}$ in numerators. In denominators, the cruder approximation $|\mathbf{r} - \mathbf{r}'| \simeq r^{(s)}$ is used.

The field emitted by the X-ray source is modeled as a sum of R independent point radiators, located at $\mathbf{r}_r''$, where $1 \leq r \leq R$. Each radiatior is assumed to emit an electric field

$$E(\mathbf{r}_r'', t'') = E_r \exp(-i\varphi_r) \exp[-i\omega_r(t'' - t_{0,r}'')] \times \theta(t'' - t_{0,r}'') \exp[-(\Delta\omega_r/2)(t'' - t_{0,r}'')]. \qquad (54)$$

In this model, each radiator has a center frequency, ω_r, and a Gaussian random amplitude and phase, $E_r \exp(-i\varphi_r)$. The start of the radiation, at $t_{0,r}''$, is accounted for by the step function, $\theta(t'' - t_{0,r}'')$ and the intrinsic line width of the radiator, $\Delta\omega_r$, is modeled as an exponential decay.

In order to calculate the field at the detector, we need the electric field at the crystal for use in equation (53):

$$E_r^{(i)}\left(\mathbf{r}', t - \frac{r^{(s)}}{c} + \frac{\mathbf{r}' \cdot \mathbf{k}^{(s)}}{ck^{(s)}}\right) = \frac{E_r \exp(-i\varphi_r)}{r^{(i)}} \exp\{-i[(\mathbf{k}_r^{(s)} - \mathbf{k}_r^{(i)}) \cdot \mathbf{r}' - \omega_r(t - t_{0,r})]\} \times \theta\left\{\frac{1}{\omega_r}[\cdot]\right\} \exp\left\{-\frac{\Delta\omega_r}{\omega_r}[\cdot]\right\}, \qquad (55)$$

where $[\cdot]$ in the last two terms denote the same argument as within the square brackets of the second exponential. We defined the relations among the coordinates of the crystal and the emitters as $-\mathbf{r}_r'' = r^{(i)}(\mathbf{k}_r^{(i)}/k_r^{(i)})$, where $\mathbf{k}_r^{(i)}$ is the wavevector of the incident wave and $k_r^{(i)} = \omega_r/c$ is its magnitude. The average retardation of the wave at the crystal is $t'' = t' - r^{(i)}/c$. The distance from the emitter to the crystal is approximated as $|\mathbf{r}' - \mathbf{r}''| \simeq r^{(i)} + \mathbf{r}' \cdot \mathbf{k}_r^{(i)}/k_r^{(i)}$ in numerators. In denominators, the cruder approximation $|\mathbf{r}' - \mathbf{r}''| \simeq r^{(i)}$ is used. The symbol $t_{0,r} = t_{0,r}'' + (r^{(i)} + r^{(s)})/c$ is the mean start time of the radiation, as seen by the detector. For elastic scattering, $k_r^{(s)} = k_r^{(i)}$.

Straightforward substitution of (55) into (53) gives

$$E^{(s)}(\mathbf{r}, t) = \frac{k_0^2}{r^{(i)}r^{(s)}} \sum_r \int_V \rho(\mathbf{r}') E_r \exp(-i\varphi_r) \times \exp\{-i[(\mathbf{k}_r^{(s)} - \mathbf{k}_r^{(i)}) \cdot \mathbf{r}' - \omega_r(t - t_{0,r})]\} \times \theta\left\{\frac{1}{\omega_r}[\cdot]\right\} \exp\left\{-\frac{\Delta\omega_r}{2\omega_r}[\cdot]\right\} \mathrm{d}\mathbf{r}'. \qquad (56)$$

We now make the approximation that the source is small and of narrow spectrum. The source then has an average frequency, ω_0, and an average wavevector, $\mathbf{k}^{(i)}$. The wavevector of the incident radiation from the rth source is $\mathbf{k}_r^{(i)} = \mathbf{k}^{(i)} + \Delta\mathbf{k}_r^{(i)}$. It is characterized by small deviations, $|\Delta\mathbf{k}_r^{(i)}| \ll |\mathbf{k}^{(i)}|$ from the average. The incident wavevector has a longitudinal component, $\mathbf{k}_{r\|}^{(i)} = \mathbf{k}^{(i)}(\boldsymbol{\omega}_r/\boldsymbol{\omega}_0) = \mathbf{k}^{(i)}(1 + \boldsymbol{\varepsilon}_r)$. It is implied that $|\boldsymbol{\varepsilon}_r| \ll 1$. To first order in ε_r, the perpendicular component, $\Delta\mathbf{k}_{r\perp}^{(i)}$, comes from the angular extent of the source only. The result, valid to first order, is then

$$\mathbf{k}_r^{(s)} - \mathbf{k}_r^{(i)} = (\mathbf{k}^{(s)} - \mathbf{k}^{(i)})(1 + \varepsilon_r) + \Delta\mathbf{k}_{r\perp}^{(i)}. \qquad (57)$$

This expression will now be substituted into equation (56) and only terms that are first order in the small quantitites ε_r, $\Delta\omega_r/\omega_r$ and $|\Delta\mathbf{k}_{r\perp}^{(i)}|/|\mathbf{k}^{(i)}|$ will be kept. The result is

$$E^{(s)}(\mathbf{r}, t) = \frac{k_0^2}{r^{(i)}r^{(s)}} \sum_r \int_V \rho(\mathbf{r}') E_r \exp(-i\varphi_r) \times \exp\{-i[(\mathbf{k}^{(s)} - \mathbf{k}^{(i)}) \cdot \mathbf{r}' - \omega_r(t - t_{0,r})]\} \times \exp\{-i[\varepsilon_r(\mathbf{k}^{(s)} - \mathbf{k}^{(i)}) \cdot \mathbf{r}' + \Delta k_{r\perp}^{(i)} \cdot \mathbf{r}']\} \times \theta\left\{\frac{1}{\omega_r}[\cdot]\right\} \exp\left\{-\frac{\Delta\omega_r}{2\omega_r}[\cdot]\right\} \mathrm{d}\mathbf{r}', \qquad (58)$$

where $[\cdot]$ in the last two terms denotes the zero-order term, $[(\mathbf{k}^{(s)} - \mathbf{k}^{(i)}) \cdot \mathbf{r}' - \omega_r(t - t_{0,r})]$.

This equation can be converted to the form

$$E^{(s)}(\mathbf{r},t)=\frac{k_0^2}{r^{(i)}r^{(s)}}\int_V \rho_{\rm equiv}(\mathbf{r}')E^{(i)}_{\rm equiv}(\mathbf{r}',t')\times\exp\{-i[(\mathbf{k}^{(s)}-\mathbf{k}^{(i)})\cdot\mathbf{r}']\}\,{\rm d}\mathbf{r}'. \quad (59)$$

The interpretation of this form is that an 'equivalent' incident X-ray field is scattered by an 'equivalent' electron density in a general direction (which is not necessarily a reciprocal-lattice vector). Comparing (58) and (59), we obtain the definition

$$\begin{aligned}\rho_{\rm equiv}(\mathbf{r}')E^{(i)}_{\rm equiv}(\mathbf{r}',t') &= \sum_r \rho(\mathbf{r}')E_r\exp(-i\varphi_r)\exp\{i[\omega_r(t-t_{0,r})]\}\\ &\times\exp\{-i[\varepsilon_r(\mathbf{k}^{(s)}-\mathbf{k}^{(i)})\cdot\mathbf{r}'+\Delta k^{(i)}_{r\perp}\cdot\mathbf{r}']\}\\ &\times\theta\left\{\frac{1}{\omega_r}[\cdot]\right\}\exp\left\{-\frac{\Delta\omega_r}{2\omega_r}[\cdot]\right\}. \quad (60)\end{aligned}$$

Here, $[\cdot]$ is defined as in (58).

In order to interpret this formula, consider a scattering point $\mathbf{r}'$ in the crystal. The electric field at that point is a sum of many random components. In particular, the factor $E_r\exp(-i\varphi_r)\exp\{i[\omega_r(t-t_{0,r})]\}$ is a Gaussian random vector for each emitter and the equivalent field is the sum over the relevant emitters. It has to be stressed that the field at any scattering point, $\mathbf{r}'$, has a well defined amplitude and phase. Let us consider now how the field changes as the scattering point moves within the crystal. The change is determined by the factor $\exp\{-i[\varepsilon_r(\mathbf{k}^{(s)}-\mathbf{k}^{(i)})\cdot\Delta\mathbf{r}'+\Delta\mathbf{k}^{(i)}_{r\perp}\cdot\Delta\mathbf{r}']\}$, where $\Delta\mathbf{r}'$ is the variation of the scattering point. In essence, if the phase of this factor changes by 2π, the new effective electric field amplitude is unrelated to the original one at $\mathbf{r}'$. We can identify the coherence volume of the incident field by the condition

$$\exp\{-i[\varepsilon_r(\mathbf{k}^{(s)}-\mathbf{k}^{(i)})\cdot\Delta\mathbf{r}'+\Delta\mathbf{k}^{(i)}_{r\perp}\cdot\Delta\mathbf{r}']\}\simeq 2\pi. \quad (61)$$

The transverse coherence length is determined by the second factor. For a small narrow-band source, it is approximately $\lambda|(\Delta\mathbf{k}^{(i)}_{r\perp}/\mathbf{k}^{(i)})|\simeq\lambda/({\rm NA})$, where (NA) is the angular extent of the X-ray source. The longitudial coherence length is determined by the relative time delay of the radiation scattered from different points in the crystal. If the difference in time delay is longer than the average coherence length of the incident radiation, the scattered waves become uncorrelated. As indicated, that happens if the first factor in (61) becomes of the order of 2π when $\Delta\mathbf{r}'$ is comparable with the unit cell. As the time delay is proportional to the momentum transfer vector, $\boldsymbol{\kappa}=\mathbf{k}^{(s)}-\mathbf{k}^{(i)}$, the scattering is always coherent at low momentum transfers, corresponding to low resolution. The diffraction becomes continuous when both the transverse and the longitudinal coherence lengths become comparable with the dimensions of the unit cell. Considering the variation of the scattering vector $\boldsymbol{\kappa}$ with scattering angle, we expect that at high resolution the diffraction pattern broadens and at even higher resolution the relevant Patterson vectors will be limited to less than a unit cell.

An approximate picture of the equivalent electron density can be given for two cases. One is an extended monochromatic source, *i.e.* Bragg scattering using a convergent incident beam. If the angular extent of the source is large enough that the transverse coherence length is comparable with the unit-cell dimension, the equivalent crystal is a row of atoms in the direction of the incident beam and the scattered intensity is the simple sum of the scattered intensities from different rows. The other case is a well collimated incident beam with finite frequency width. This could be categorized as narrow-band Laue diffraction. The 'longitudinal' coherence length is now in the direction of the scattering vector, characterized by (H,K,L) in reciprocal space. We obtain incoherence between two unit cells when the fractional bandwidth of the incident radiation is of the order of the resolution (in the proper units). It can be expressed simply as $|\boldsymbol{\varepsilon}_r|\cdot|(H,K,L)|\simeq 1$. At low scattering angles, the equivalent crystal is a slab perpendicular to the scattering vector. If the slab is thick enough, we recover the usual picture of the Laue method, where the crystal selects the proper scattering wavelength. As the scattering angle increases, the equivalent crystal becomes thinner and thinner. We expect, therefore, both a change in scattering angle (Dušek, 1994) and a decrease in wavelength selectivity.

Finally, the connection with the mutual-coherence function can be made by substituting our radiation model, equation (54), into expression (17) and finally evaluating equation (20). The calculation is straightforward. As previously, we use scalar fields and complex notation. The mutual-correlation function has 'diagonal' terms, in which the radiators are considered one at a time and 'off-diagonal' terms where the different radiators interfere. Only the diagonal terms will be presented here. The off-diagonal terms are, in general, of comparable magnitude. (In the case of almost monochromatic radiators, with a relatively large frequency spread, the diagonal terms dominate.) The result is

$$\Gamma^{(i)}(\mathbf{r}';\mathbf{r}'';T)=\frac{1}{(r^{(i)})^2}\sum_r\frac{E_r^2}{\Delta\omega_r}\exp\{-i[\omega_rT-\mathbf{k}^{(i)}_r\cdot(\mathbf{r}''-\mathbf{r}')]\}\times\exp\left\{-\frac{\Delta\omega_r}{2\omega_r}[\cdot]\right\}. \quad (62)$$

Using the same assumptions that led to equations (60) and (61), we obtain

$$\begin{aligned}\Gamma^{(i)}(\mathbf{r}';\mathbf{r}'';\Delta\mathbf{r}'\cdot\hat{\mathbf{s}}/c) &= \exp\{-i[(\mathbf{k}^{(s)}-\mathbf{k}^{(i)})\cdot\Delta\mathbf{r}']\}\frac{1}{(r^{(i)})^2}\sum_r\frac{E_r^2}{\Delta\omega_r}\\ &\times\exp\{-i[\varepsilon_r(\mathbf{k}^{(s)}-\mathbf{k}^{(i)})\cdot\Delta\mathbf{r}'+\Delta k^{(i)}_{r\perp}\cdot\Delta\mathbf{r}']\}\\ &\times\exp\left[-\frac{\Delta\omega_r}{2\omega_r}(\mathbf{k}^{(s)}-\mathbf{k}^{(i)})\cdot\Delta\mathbf{r}'\right]. \quad (63)\end{aligned}$$

A comparison of (63) with (60), (61) and (20) reveals the marked similarity of all our results.

I would like to start by acknowledging David Sayre's contributions to the subject matter of this paper. He was by far the first to understand the equivalence of the crystallographic phase problem and the sampling theorem of information theory. I would like to thank him for very helpful discussions

over the years. Well deserved thanks are given to my host in Uppsala, János Hajdú, for his support, encouragement and discussions on the subject. My thanks are also gladly extended to Edgar Weckert and Henry Chapman for forcing me to rethink the calculations in terms of electric fields and making the results more realistic. Finally, I would like to thank Hanna Szöke for carefully reading the manuscript. I submit that all the matter in this paper was developed independently and that I have done my best to credit every source. Part of this work was performed under the auspices of the US Department of Energy, under contract No. W-7405-ENG.-48, and part of it was supported by the Swedish Strategic Foundation, STINT.

References

Barton, J. J. (1991). *Phys. Rev. Lett.* **67**, 3106–3109.
Bates, R. H. T. (1982). *Optik* (*Stuttgart*), **61**, 247–262.
Béran, P. & Szöke, A. (1995). *Acta Cryst.* A**51**, 20–27.
Berne, B. J. & Pecora, R. (1976). *Dynamic Light Scattering: with Applications to Chemistry, Biology, and Physics.* New York: Wiley.
Bertero, M. (1989). *Linear Inverse and Ill Posed Problems. Advances in Electronics and Electron Physics*, Vol. 75. New York: Academic Press.
Born, M. & Wolf, E. (1980). *Principles of Optics*, 6th ed. Oxford: Pergamon.
Bricogne, G. (1992). In *International Tables for Crystallography*. Vol. B, edited by U. Shmueli. Dordrecht: Kluwer Academic Publishers.
Cowley, J. M. (1995). *Diffraction Physics*, 3rd ed. Amsterdam: North-Holland.
Daubechies, I. (1992). *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM.
Dušek, M. (1994). *Optics Commun.* **111**, 203–208.
Dušek, M. (1995). *Phys. Rev. E*, **52**, 6833–6840.
Fetter, A. L. & Walecka, J. D. (1971). *Quantum Theory of Many Particle Systems.* New York: McGraw-Hill.
Gábor, D. (1948). *Nature (London)*, **161**, 777–778.
Gábor, D. (1949). *Proc. R. Soc. (London) Ser. A*, **197**, 454–487.
Gog, T., Len, M., Materlik, G., Bahr, D., Fadley, C. S. & Sanchez-Hanke, C. (1996). *Phys. Rev. Lett.* **76**, 3132–3136.
Goodman, J. W. (1985). *Statistical Optics*. New York: Wiley.
Jackson, J. D. (1975). *Classical Electrodynamics*, 2nd ed. New York: Wiley.
Jagodzinski, H. & Frey, F. (1992). In *International Tables for Crystallography*. Vol. B, edited by U. Shmueli. Dordrecht: Kluwer Academic Publishers.
James, R. W. (1982). *The Optical Principles of the Diffraction of X-rays*. Reprinted by Ox Bow Press, Woodbridge.
Kampen, N. G. van (1969). *Proceedings of the International School of Physics*, *Enrico Fermi*, *Course XLII*, edited by R. J. Glauber. New York: Academic Press.
Leith, E. N. & Upatnieks, J. (1962). *J. Opt. Soc. Am.* **52**, 1123–1130.
Luenberger, D. G. (1984). *Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison Wesley.
Mandel, L. (1969). *Phys. Rev.* **181**, 75–84.
Mandel, L. & Wolf, E. (1995). *Optical Coherence and Quantum Optics*. New York: Cambridge University Press.
Miao, J., Charalambous, P., Kirz, J. & Sayre, D. (1999). *Nature (London)*, **400**, 342–344.
Miao, J. & Sayre, D. (2000). *Acta Cryst.* A**56**, 596–605.
Miao, J., Sayre, D. & Chapman, H. N. (1998). *J. Opt. Soc. Am.* **15**, 1662–1669.
Natterer, F. (1986). *The Mathematics of Computerized Tomography*. Chichester: Wiley.
Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature (London)*, **406**, 752–757.
Newton, R. G. (1982). *Scattering Theory of Waves and Particles*, 2nd ed. New York: Springer-Verlag.
Papoulis, A. (1991). *Probability, Random Variables and Stochastic Processes*. New York: McGraw Hill.
Pendry, J. B. (1974). *Low Energy Electron Diffraction.* London: Academic Press.
Sabatier, P. C. (1987). Editor. *Basic Methods of Tomography and Inverse Problems*. London: Adam Hilger.
Saldin, D. K. (1997). *Surf. Rev. Lett.* **4**, 441–457.
Saldin, D. K. & de Andres, P. L. (1990). *Phys. Rev. Lett.* **64**, 1270–1274.
Sayre, D. (1952). *Acta Cryst.* **5**, 843–843.
Shannon, C. E. (1949). *Prooc. IRE*, **37**, 10–21.
Siegman, A. E. (1986). *Lasers*. Mill Valley, CA: University Science.
Stark, H. (1987). Editor. *Image Recovery: Theory and Application*. Orlando: Academic Press.
Szöke, A. (1986). *Short Wavelength Coherent Radiation: Generation and Applications*. *AIP Conference Proceedings*, No. 147, edited by D. T. Attwood & J. Bokor. New York: American Institute of Physics.
Szöke, A. (1993). *Acta Cryst.* A**49**, 853–866.
Szöke, A. (1997). *J. Imag. Sci. Technol.* **41**, 332–341.
Szöke, A. (1998). *Acta Cryst.* A**54**, 543–562.
Szöke, A. (1999). *Chem. Phys. Lett.* **313**, 777–788.
Szöke, A., Szöke, H. & Somoza, J. R. (1997). *Acta Cryst.* A**53**, 291–313.
Tegze, M. & Faigel, Gy. (1996). *Nature (London)*, **380**, 49–52.
Tegze, M., Faigel, Gy., Marchesini, S., Belakhovsky, M. & Ulrich, O. (2000). *Nature (London)*, **407**, 38–38.
Tollin, P., Main, P., Rossmann, M. G., Stroke, G. W. & Restrick, R. C. (1966). *Nature (London)*, **209**, 603–604.
Weckert, E. & Hümmer, K. (1997). *Acta Cryst.* A**53**, 108–143.
Wolf, E. & Foley, J. T. (1989). *Phys. Rev. A*, **40**, 579–587.
Woolfson, M. & Fan, H.-F. (1995). *Physical and Non-Physical Methods of Solving Crystal Structures*. Cambridge University Press.