Partially coherent image formation with x-ray microscopes

L. Jochum and W. Meyer-Ilse

Image formation with partially coherent radiation is evaluated with the Hopkins formula and then applied to x-ray microscopy. Image characteristics expected from instruments with circular and annular pupils in partially coherent conditions are considered for two-point objects and a knife-edge object. The theoretically expected values for image characteristics that are easy accessible by an experiment, such as the width of a knife edge, are given for various x-ray microscopes.

Keywords: Image formation, partially coherent imaging, x-ray microscopy, resolution limit.

1. Introduction

X-ray microscopy is a unique technique that extends visible light microscopy to higher resolution and makes use of unique contrast mechanisms. It does not compete with techniques such as electron microscopy in terms of resolution but rather offers unique advantages including the ability to image samples in an aqueous environment. X-ray microscopy yields information with high resolution from thick samples that cannot be obtained by any other technique including methods such as atomic force microscopy or near-field microscopy. Applications of x-ray microscopy are in several fields including biology and material science. The interaction of x rays with matter provides unique elemental and chemical contrast. Construction of an x-ray microscopy resource center is planned at the Advanced Light Source ALS, in Berkeley, Calif., that will provide a broad user community with the necessary instrumentation to use this technique optimally.

There are two basic types of x-ray microscope, conventional x-ray microscopes (XM's) and scanning x-ray microscopes (SXM's). XM's are similar to their visible light counterparts with critical illumination. Sometimes they are also called imaging or full-field x-ray microscopes. XM's with zone plates were pioneered at the University of Gottingen. The most prominent instruments of each kind are the Göttingen XM at the Berlin electron storage ring (BESSY) and the STXM operated at the Brookhaven National Laboratory (BNL). Because these two microscope types are complementary, there are plans to build them at the ALS. An XM that uses bending magnet radiation is completed and has been operational since the Fall of 1994.

The resolution of x-ray microscopes is currently not limited by the wavelength of the imaging light but rather by the ability to fabricate optical elements that provide high-resolution images. The highest resolution for both types of microscope is achieved with zone plate lenses. So far, the numerical aperture of the zone plates is typically below NA = 0.1. For microscopes operating in the water window (2.4 nm ≤ λ ≤ 4.5 nm), a resolution approximately five times better than with visible light is commonly achieved.

Unlike conventional refractive lenses, the focusing properties of zone plates are based on diffraction. This implies that the focal length of a zone plate lens strongly depends on the wavelength and that multiple diffraction orders exist. In most cases only the first diffraction order is used. The diffraction efficiency in the first diffraction order of objective zone plates is currently in the range of 10%, with improvements expected in the future. If operated with quasi-monochromatic radiation, and if the number of zones is large, zone plate lenses obey the same laws as refractive lenses.

In the scope of this paper we assume that the above-mentioned assumptions are fulfilled and treat zone plates as conventional lenses. Therefore the conclusions in this paper also hold for microscopes other than x-ray microscopes.

2. X-Ray Microscopes

A schematic of the optical layout of an XM like the one at BESSY and the one at the ALS is shown in Fig.
Polychromatic synchrotron radiation is focused onto the sample by a condenser zone plate, and an objective zone plate forms an enlarged high-resolution image of the object. This image is detected with an x-ray CCD camera. Because zone plates are diffraction optics, they have a strong chromatic error, and the image quality depends on the monochromaticity of the image-forming light. To provide the required monochromaticity, a pinhole is put in close proximity to the sample plane. This pinhole together with the condenser zone plate acts as a linear monochromator. With \( D \) as the diameter of the zone plate and \( d \) as the pinhole diameter, the spectral bandwidth of the illuminating light is given by \( \Delta \lambda / \lambda = 2d/D \). The center wavelength can be tuned by changing the distance between the condenser zone plate and the pinhole.

A schematic of an SXM like the one at the BNL is shown in Fig. 1b. A high-resolution objective zone plate is spatially coherent illuminated and forms a diffraction-limited x-ray spot. This spot is scanned over the sample area while the transmitted x-ray intensity is recorded by a detector. Because only the spatially coherent fraction of the x-ray source can be used, such an instrument would preferably be built at an undulator light source to achieve reasonably short exposure times. A central stop on the objective zone plate and an order-sorting aperture between the objective zone plate and the sample are used to block out zero-order radiation.

The main advantage of an SXM compared with an XM is the absence of an objective zone plate downstream of the sample. The low efficiency of the objective zone plate contributes to the transmission losses in an XM. Therefore an SXM requires a smaller radiation dose on the sample to obtain the same signal-to-noise ratio as an XM.

To take advantage of the increased resolution compared with visible light microscopy, microscopy with x-ray-excited visible light luminescence requires an SXM.

3. Image Formation in X-Ray Microscopes

Microscopes use a probe to determine the distribution of one or more physical quantities in the sample with high spatial resolution. In x-ray microscopy the probe consists of x rays. For low-energy x rays as used in x-ray microscopy, the main interactions are photoelectric absorption and elastic scattering, whereas inelastic scattering (Compton scattering) may be neglected. In a single-atom model, the interaction of any material with x rays can be described by the complex atomic scattering factor \( f_1 = f_1 + if_2 \); values of \( f_1 \) and \( f_2 \) are listed in Ref. 10. On a macroscopic scale the interaction of x rays with matter is characterized by the complex index of refraction \( \eta = 1 - \lambda / \lambda_0 - i \lambda / 4\pi \mu \), where \( \mu \) is the linear absorption coefficient and \( \eta \) is the phase shift coefficient of the wave with respect to its propagation in vacuum. The atomic scattering factors are related to \( \mu \) and \( \eta \) as \( \mu = 2r_0\lambda\eta_\lambda f_1 \) and \( \eta = r_0\lambda\eta_\lambda f_2 \) with \( r_0 \) being the classical electron radius and \( \eta_\lambda \) the number of atoms per unit volume. In x-ray microscopes the sample is imaged in transmission. This image contains the information about the spatial distribution of the complex index of refraction \( \eta x, y, z \) in the sample. For the low-energy x rays used in x-ray microscopy, where inelastic scattering processes can be neglected, the image formation for both types of x-ray microscope (SXM and XM) may be described by the same formalism. The equivalence of scanning and conventional microscopes is valid even for imaging thick objects. The differences between the two instruments are purely technical. Therefore the conclusions and results in this paper apply for both types of x-ray microscope, although the
parameters below have different meanings, depending on the type of instrument.

The microscopes considered here detect the intensity distribution in the detector plane, so that only one two-dimensional 2D section of the sample can be seen at a time. To record the complex index of refraction \( n(x, y, z) \) in all three spatial dimensions with such an instrument, one needs to take several images with different views of the sample. To gather three-dimensional 3D information in one exposure by using only one detector plane, the recording of the amplitude and the phase of the wave transmitted by a 3D object cannot be distinguished from a 2D object that produces the identical wave field in a plane downstream of the object. Below we briefly review the theory of partially coherent image formation and with that background describe the properties of a 2D view of different samples in an x-ray microscope.

4. Partially Coherent Imaging

The theory of partially coherent image formation was first developed by Hopkins.\(^{1,14}\) If \( x_0, y_0 \) and \( x_1, y_1 \) are Cartesian coordinates in the object plane and in the image plane, respectively, the intensity profile \( I(x_1, y_1) \) in the image plane of an isoplanatic microscope is

\[
I(x_1, y_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(x_0 - x'_0, y_0 - y'_0) \times F[x_0, y_0] F^*[x'_0, y'_0] K(x_1 - x_0, y_1 - y_0) \times K^*[x'_1 - x'_0, y'_1 - y'_0] dx_0 dy_0 dx'_0 dy'_0.
\]

In Eq. \( I \) \( J(x, y) \) is the mutual intensity in the object plane, \( F \) \( F^* \) is the complex transmission of the object, and \( K \) \( K^* \) is the Fourier transform of the complex transmission of the objective lens. Any objective lens aberrations or effects of defocusing can be included in \( K \).

The physical meaning of the mutual intensity is that it characterizes the phase correlation between any two points in the object plane. If the illuminating light is perfectly incoherent, the phases vary statistically so that the mutual intensity equals zero unless the two points are identical. Mathematically, this behavior is described by the Dirac \( \delta \) function, and Eq. \( I \) reduces to the well-known linear relation between the object and the image intensity with the intensity point-spread function as the linear filter. For totally coherent illumination the phase correlation and thus the mutual intensity is constant for any two object points. In this case Eq. \( I \) describes a linear transformation of the complex amplitude from the object to the image plane with the amplitude point-spread function as the linear filter. To make use of this linearity, the detector must record the amplitude and phase in the image plane as it is done in holography.

The results of investigations of the resolution in x-ray microscopy for the limiting cases of coherent and incoherent imaging have been published in Ref. 15. In general, the image formation in a microscope is partially coherent. As we show below, the effects of partial coherence are most important for feature sizes near the limit of resolution.

When imaging with a circular symmetric aberration-free objective lens that might have a central stop of relative radius \( 0 < \varepsilon_k < 1 \), the objective lens function \( K(x, y) \) is given by

\[
K(x, y) = \frac{2}{1 - \varepsilon_k^2} \left| J_{1/2}(\varepsilon_k m v) \right| \frac{\varepsilon_{1/2}(\varepsilon_k m v)}{m v}
\]

with \( v = 2\pi x^2 + y^2 \).\(^1\)

\[
J_{1/2}(\varepsilon_k m v) = \frac{2}{1 - \varepsilon_k^2} \left| J_{1/2}(\varepsilon_k m v) \right|
\]

with \( v = 2\pi x^2 + y^2 \).\(^3\)

The physical meaning of the mutual intensity is that it characterizes the phase correlation between any two points in the object plane. If the illuminating light is perfectly incoherent, the phases vary statistically so that the mutual intensity equals zero unless the two points are identical. Mathematically, this behavior is described by the Dirac \( \delta \) function, and Eq. \( I \) reduces to the well-known linear relation between the object and the image intensity with the intensity point-spread function as the linear filter. For totally coherent illumination the phase correlation and thus the mutual intensity is constant for any two object points. In this case Eq. \( I \) describes a linear transformation of the complex amplitude from the object to the image plane with the amplitude point-spread function as the linear filter. To make use of this linearity, the detector must record the amplitude and phase in the image plane as it is done in holography.

The results of investigations of the resolution in x-ray microscopy for the limiting cases of coherent and incoherent imaging have been published in Ref. 15. In general, the image formation in a microscope is partially coherent. As we show below, the effects of partial coherence are most important for feature sizes near the limit of resolution.

When imaging with a circular symmetric aberration-free objective lens that might have a central stop of relative radius \( 0 < \varepsilon_k < 1 \), the objective lens function \( K(x, y) \) is given by

\[
K(x, y) = \frac{2}{1 - \varepsilon_k^2} \left| J_{1/2}(\varepsilon_k m v) \right| \frac{\varepsilon_{1/2}(\varepsilon_k m v)}{m v}
\]

with \( v = 2\pi x^2 + y^2 \).\(^1\)

\[
J_{1/2}(\varepsilon_k m v) = \frac{2}{1 - \varepsilon_k^2} \left| J_{1/2}(\varepsilon_k m v) \right|
\]

with \( v = 2\pi x^2 + y^2 \).\(^3\)

The physical meaning of the mutual intensity is that it characterizes the phase correlation between any two points in the object plane. If the illuminating light is perfectly incoherent, the phases vary statistically so that the mutual intensity equals zero unless the two points are identical. Mathematically, this behavior is described by the Dirac \( \delta \) function, and Eq. \( I \) reduces to the well-known linear relation between the object and the image intensity with the intensity point-spread function as the linear filter. For totally coherent illumination the phase correlation and thus the mutual intensity is constant for any two object points. In this case Eq. \( I \) describes a linear transformation of the complex amplitude from the object to the image plane with the amplitude point-spread function as the linear filter. To make use of this linearity, the detector must record the amplitude and phase in the image plane as it is done in holography.

The physical meaning of the mutual intensity is that it characterizes the phase correlation between any two points in the object plane. If the illuminating light is perfectly incoherent, the phases vary statistically so that the mutual intensity equals zero unless the two points are identical. Mathematically, this behavior is described by the Dirac \( \delta \) function, and Eq. \( I \) reduces to the well-known linear relation between the object and the image intensity with the intensity point-spread function as the linear filter. For totally coherent illumination the phase correlation and thus the mutual intensity is constant for any two object points. In this case Eq. \( I \) describes a linear transformation of the complex amplitude from the object to the image plane with the amplitude point-spread function as the linear filter. To make use of this linearity, the detector must record the amplitude and phase in the image plane as it is done in holography.

The physical meaning of the mutual intensity is that it characterizes the phase correlation between any two points in the object plane. If the illuminating light is perfectly incoherent, the phases vary statistically so that the mutual intensity equals zero unless the two points are identical. Mathematically, this behavior is described by the Dirac \( \delta \) function, and Eq. \( I \) reduces to the well-known linear relation between the object and the image intensity with the intensity point-spread function as the linear filter. For totally coherent illumination the phase correlation and thus the mutual intensity is constant for any two object points. In this case Eq. \( I \) describes a linear transformation of the complex amplitude from the object to the image plane with the amplitude point-spread function as the linear filter. To make use of this linearity, the detector must record the amplitude and phase in the image plane as it is done in holography.
with \( \lambda_c \) and \( \lambda_o \) being the wavelengths that pass the condenser and objective lens and \( \text{NA}_c \) and \( \text{NA}_o \) being the numerical apertures of the condenser and objective lens, respectively. As for soft x rays, where inelastic scattering can be neglected, \( \lambda_c \) equals \( \lambda_o \), the coherence parameter \( m \) is given by the ratio of the numerical apertures of the condenser and the objective lens. Equation (4) also holds for an SXM if \( \text{NA}_c \) is replaced by the sine of the half angle of the acceptance cone of a circular detector.

If either the object itself or its Fourier spectrum is javascript:off; \( \epsilon_j \) or \( \epsilon_k \) is nonzero, the image contrast of a two-point object imaged with a contrast of more than 40% with a partially coherent imaging system and an appropriate coherence parameter \( m \). For a fixed point distance \( d \) and increasing \( m \), the image contrast steadily increases to a maximum between \( m = 1 \) and \( m = 1.5 \), which is shown in Fig. 3. As \( m \) approaches infinity, the contrast is slightly oscillating around the contrast obtained with an incoherent imaging system \( (m \to \infty) \).

Figure 4 and 5 show the influence of different central stops on the image contrast. The coherence parameter is set to \( m = 1 \) for these examples. The mutual intensity \( J(x, y) \) in the object plane varies with the parameter \( \epsilon_j \) whereas \( \epsilon_k \) describes a central stop on the objective lens and thus modifies \( K(x, y) \). For a two-point object a central stop on the objective lens \( (\epsilon_k > 0) \) has a greater effect on the image contrast than a change of the mutual intensity from a similar stop on the condenser lens \( (\epsilon_j > 0) \).

Fig. 2. Image contrast of a two-point object imaged with \( \epsilon_j = \epsilon_k = 0 \) and coherence parameters \( m \) of 0, 0.5, 1, and infinity; \( d \) is in units of \( \lambda / \text{NA}_c \).

Fig. 3. Influence of \( m \) on the image contrast of a two-point object for four selected point distances \( d \) imaged with circular pupils \( (\epsilon_j = \epsilon_k = 0; d \text{ in units of } \lambda / \text{NA}_o) \).

Fig. 4. Image contrast of a two-point object versus the inverse of the point distance \( 1/d \) in units of \( \lambda / \text{NA}_c \) for \( m = 1 \) and \( a. \epsilon_k = 0 \) for different mutual intensities \( 0 \leq \epsilon_j \leq 0.9 \). (b) \( \epsilon_j = 0 \) for different central stops on the objective lens \( 0 \leq \epsilon_k \leq 0.9 \).
For any point distance \(d\) an increase in \(\varepsilon_k\) results in a higher image contrast. The contrast enhancement by changing \(\varepsilon_k\) is less significant for larger point distances, which can be seen as different slopes of the curves in Fig. 5b.

If \(m = 1\) and \(\varepsilon_j = \varepsilon_k = 0\), a point separation of \(d = 0.61\) corresponds to the resolution limit according to the Rayleigh criterion [Ref. 14, p. 333]. Applying the Rayleigh criterion to an incoherent imaging system with circular pupils leads to an image contrast of 15.3% for a two-point object. By changing \(\varepsilon_k\) from 0 to 0.9 at a point distance of \(d = 0.61\), the image contrast increases from 15.3% to almost 70% [see Fig. 5b].

Let \(L\) be the distance of two object points at which the image contrast is 15.3%. In unnormalized coordinates this distance is \(6x = Lx/\lambda NA_0\). Figure 6 shows \(L\) as a function of \(m\) for various \(\varepsilon_j\) and \(\varepsilon_k\). In the range of \(0 \leq m \leq 1\), which is relevant for x-ray microscopy, an increase of \(\varepsilon_j\) as well as \(\varepsilon_k\) results in a decreasing \(L\). This can be interpreted as an improved resolution, but one must remember that these results refer only to the imaging of a two-point object.

6. Partially Coherent Imaging of a Knife Edge

After examining the behavior of a partially coherent imaging system in the case of a two-point object, we investigated the same system with a 2D knife edge as an object. Analyzing the image of a knife edge is a very common and convenient method to check the image quality experimentally. To simulate the partially coherent imaging of a knife edge, Eq. 1 was evaluated numerically with the program SPLAT, developed by Toh at the Department of Electrical Engineering and Computer Sciences, University of California.

Figure 7 shows two examples of an edge response and the differentiated response. The imaging parameters in Fig. 7a represent a typical layout for the XM-1 at the ALS, and in Fig. 6b are typical parameters for an SXM.

The full width at half-maximum (FWHM) of the differentiated image from an edge is often used to determine the resolution of a system experimentally.
Table 1. Image Characteristics of a Knife-Edge Object and a Two-Point Object for Different Optical Setups

<table>
<thead>
<tr>
<th>Optical Setup</th>
<th>m</th>
<th>$\epsilon_j$</th>
<th>$\epsilon_k$</th>
<th>Different Edge Response FWHM ($\lambda / NA$)</th>
<th>Edge Response 25–75% ($\lambda / NA$)</th>
<th>Edge Response 10–90% ($\lambda / NA$)</th>
<th>L = Two-Point Distance for 15.3% Contrast</th>
<th>Two-Point Contrast at $d = 0.61 %$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incoherent</td>
<td>$\infty$</td>
<td>0</td>
<td>0</td>
<td>0.49</td>
<td>0.31</td>
<td>0.62</td>
<td>0.61</td>
<td>15.3</td>
</tr>
<tr>
<td>SXM-lum</td>
<td>$\infty$</td>
<td>0</td>
<td>0.45</td>
<td>0.49</td>
<td>0.31</td>
<td>0.62</td>
<td>0.61</td>
<td>15.3</td>
</tr>
<tr>
<td>SXM</td>
<td>1</td>
<td>0</td>
<td>0.45</td>
<td>0.45</td>
<td>0.40</td>
<td>1.31</td>
<td>0.57</td>
<td>29.5</td>
</tr>
<tr>
<td>XM-2</td>
<td>1</td>
<td>0</td>
<td>0.45</td>
<td>0.47</td>
<td>0.42</td>
<td>1.32</td>
<td>0.55</td>
<td>29.5</td>
</tr>
<tr>
<td>XM-1, 30 nm</td>
<td>0.54</td>
<td>0.33</td>
<td>0</td>
<td>0.47</td>
<td>0.29</td>
<td>0.59</td>
<td>0.61</td>
<td>15.3</td>
</tr>
<tr>
<td>XM-1, 25 nm</td>
<td>0.45</td>
<td>0.33</td>
<td>0</td>
<td>0.45</td>
<td>0.22</td>
<td>0.39</td>
<td>0.70</td>
<td>0</td>
</tr>
<tr>
<td>Coherent</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.42</td>
<td>0.21</td>
<td>0.37</td>
<td>0.73</td>
<td>0</td>
</tr>
</tbody>
</table>

The indicated value for the FWHM in Fig. 7 is what can be theoretically expected from that particular instrumental configuration. FWHM values of the differentiated edge response for different instrument configurations are tabulated in Table 1. Although the intensity profiles of the edge response in Fig. 7 look quite different, they result in the same FWHM of the differentiated response. This is because the FWHM, like the image contrast of a two-point object, is mainly characterized by the width of the central maximum of $K_x y$. By introducing a central stop in the objective lens, the central maximum of $K_x y$ becomes narrower, but the higher-order fringes around the maximum increase in intensity, as is shown in Fig. 8. As previously shown, the narrower central maximum results in a higher image contrast in the case of a two-point object, especially for small point distances. When a knife edge is imaged with the same optical system, the increased intensity in the higher-order fringes blur the image, as seen in Fig. 7b, and tend to reduce the overall contrast of an arbitrary object. When looking at the distance of a 90–10% intensity drop in the edge response, which is the corresponding quantity to the Rayleigh resolution limit of an incoherent imaging system, we see that this distance is significantly larger in Fig. 7b than in Fig. 7a. Other then the FWHM, which is the same for both cases, this distance is sensitive to the changed parameters of the optical system and thus is a more likely candidate to characterize the quality and performance of a particular imaging system.

7. Summary

The properties of the image formation with partially coherent light have been described for x-ray microscopes. The results are not restricted to x-ray microscopy; they also apply to other microscopes including visible light and transmission electron microscopes. The results of our evaluations are summarized in Table 1. It shows the image characteristics of a knife-edge object and a two-point object for different combinations of $m$, $\epsilon_j$, and $\epsilon_k$. Each row represents a particular optical setup. The first and the last rows show the theoretical cases of a totally incoherent and a totally coherent system, respectively. All the other optical setups shown in Table 1 are likely candidates for the microscopes planned at the ALS.

SXM-lum stands for a scanning x-ray microscope operating with the x-ray excitation of visible light luminescence. Because the luminescence light is incoherent, the coherence parameter $m$ is set to infinity. An unobstructed detector and a central stop on the objective zone plate define the parameters $\epsilon_j$ and $\epsilon_k$, respectively. The incoherent image formation implies that the image formation is linear in intensity, even beyond the diffraction limit of resolution. This enables techniques that use inverse filtering to overcome the diffraction limit (Ref. 17, p. 133ff).

The SXM in Table 1 is the same instrument but not in the luminescence mode, so that $m$ is given by the numerical aperture of the objective zone plate and half of the acceptance angle of the detector. XM-2 is a conventional microscope. The SXM and the XM-2 are to operate at a short-period undulator source. The optical design of XM-2 will be such that it has a circular entrance and exit pupils and matching numerical apertures of the illuminating and imaging optics. XM-1 is the conventional microscope that operates with bending magnet radiation. In Table 1 its imaging properties are shown for two different objective zone plates, one with a 30-nm and one with a 25-nm outermost zone width. In both cases it has a large diameter condenser zone plate with a 55-nm outermost zone width, so that these two setups have a different coherence parameter $m$.

Comparing the first two rows in Table 1 or the SXM and the XM-2 cases, one can see how an annular exit pupil improves the two-point resolution but increases...
the width of the edge response at the same time. For circular exit pupils \( \epsilon_c = 0 \), the edge width decreases so that it appears sharper with a decreasing coherence parameter \( m \), whereas the two-point resolution is reduced at the same time.

This object dependence of the image formation is caused by the nonlinearity of Eq. 1. The fact that knife-edge images produced with partially coherent radiation appear sharper compared with the incoherent case is important for the characterization of instruments near their diffraction limit. Looking at the numbers in the first three rows of Table 1, one can see that a central stop on the objective zone plate, which is necessary in a SXM, affects the image characteristics significantly compared with the influence of the coherence parameter. In a knife-edge test an instrument such as the SXMinTable 1 should produce a 75–25% image intensity drop at a distance of 0.4 \( \lambda / NA \) or a FWHM of 0.45 \( \lambda / NA \), to be considered as diffraction limited.

We thank Derek Leed from the Department of Electrical Engineering and Computer Sciences, University of California, for helpful support with the program SPLAT, which was used to calculate the edge response data. This research is supported by the United States Department of Energy, Office of Basic Energy Sciences, and the Office of Health and Environmental Research under contract DE-AC03-76SF00098.

References


