PHYSICS OF OUR DAYS

Contents

Wave packet in the phase problem in optics and ptychography

N L Popov, I A Artyukov, A V Vinogradov, V V Protopopov

DOI: https://doi.org/10.3367/UFNe.2020.05.038775

1.	Introduction	766
2.	Retrieval of the coherent field phase and modulus. The wave packet method	767
3.	Phase retrieval discretization in the wave packet method	768
4.	Results of calculations. Comparison of the wave packet and paraxial approximation methods	770
5.	Ptychography. Results of calculations	771
6.	Conclusions	773
	References	773

Abstract. At present, ptychography seems to be the most natural and efficient method for approaching the diffraction-limited optical resolution. The general setup of a ptychoscope does not contain refracting or focusing elements and includes a coherent illumination source, a translation stage for displacement of a macroscopic object, and a detector for recording transmitted or reflected radiation from the object, which is connected to a computer for processing diffractograms. In classical optics, the main problem with achieving high spatial resolution is the correction and elimination of aberrations in optical systems, whereas the spatial resolution in ptychography mainly depends on the reliability of recording and computer processing diffractograms with large numerical apertures. After a brief introduction to the history and current state of ptychography, the wave-packet method for calculating the wave field on a detector in the far field and for a large numerical aperture is considered in detail. This gives a relation between fields on the object and on the detector, which underlies the ePIE (extended Ptychography Iterative Engine) algorithms for recovering images used in practice. The realization of algorithms involves operations with functions defined in certain domains (coordinate networks) of the direct space and Fourier space related to the object and detector. The size of and steps involved in such networks are strictly related to the object size, its distance from the detector, and the numerical aperture. The programs developed in this paper are used to refine the limits of applicability of the paraxial approximation (Fresnel integrals) in calculations of the field on the detector. Simulations of images obtained by the ptychography method are presented.

Keywords: ptychography, phase retrieval

N L Popov⁽¹⁾, I A Artyukov⁽²⁾, A V Vinogradov⁽³⁾, V V Protopopov Lebedev Physical Institute, Russian Academy of Sciences, Leninskii prosp. 53, 119991 Moscow, Russian Federation E-mail: ⁽¹⁾ popovnl@sci.lebedev.ru, ⁽²⁾ iart@mail.ru, ⁽³⁾ vinograd@sci.lebedev.ru

Received 19 January 2020, revised 12 May 2020 Uspekhi Fizicheskikh Nauk **190** (8) 820–828 (2020) Translated by M Sapozhnikov; edited by A M Semikhatov

1. Introduction

The term 'phase problem in optics' is closely related to the concept of 'lensless optics'. In both cases, a problem arises that was initially formulated as the retrieval of the phase of a wave field produced by a laser beam by measuring its modulus in two parallel planes: of the object and of the detector. Iterative calculations of the propagation of coherent radiation from the object plane to the detector plane are used to obtain distributions of the complex field amplitude in both planes. A similar problem was first considered back in 1972 [10] using one of the best computers at that time with a memory of ~ 1 MB and a rate of ~ 1 megaflop, providing the analysis of numerous 32×32 images. The development of the ideas proposed in that paper resulted in the appearance of lensless imaging¹ in the form used at present in the visible, vacuum UV, an X-ray ranges [2–9].² The initial algorithm [1] has been modified many times, becoming more efficient, in particular, due to the improvement in computers. Algorithms were developed based on the use of a priori information instead of measurements of the intensity distribution on the object. As a result, lensless imaging methods are now used in physical studies, and related commercial projects have appeared [17–20]. Currently the most popular among them is *ptychography* [21, 22], the method for the retrieval of the field amplitude and phase on an object by computer processing of its diffractograms during the subsequent known displacement of the object with respect to the illuminating beam and the detector. In this case, adjacent fields of view should overlap by no less than a half. Ptychography is described in more detail in Section 5.

One of the main advantages of lensless imaging and ptychography is the fundamental possibility of bypassing problems related to aberrations of optical elements (lenses, mirrors, zone plates, etc.) and approaching the diffractionlimited resolution in microscopy and atmospheric and astronomical optics. Obviously, to fully realize this advantage of lensless systems, it is necessary to think outside the

¹ Imaging without using optical elements placed between an object and the detector of radiation from the object.

² Examples of work done in this country are [10-16].

scope of paraxial methods of simulating the propagation of light beams from the object to the detector. This paper is devoted to solving this problem within the framework of the scalar wave equation.

2. Retrieval of the coherent field phase and modulus. The wave packet method

Most of the modern phase-retrieval methods are essentially the development of algorithms formulated in work performed 30-40 years ago (see the references in [10, 23, 24]). Algorithms presented in many papers and monographs are still being refined. These algorithms allow finding a complex field (intensity and phase) on the object surface from the measured field intensity on the detector. We recall that optical elements between the object and detector are absent; however, a priori information on the object properties is used. As a side result of the algorithm operation, the phase distribution on the detector surface is obtained. In practice, the most convenient diffraction integral for a given experiment is used, which expresses the field on the detector in terms of the field distribution on the object surface. As a rule, the calculation of the diffraction integral reduces to the calculation of direct and inverse Fourier transforms.

The algorithms are based on the *multiple* application of four main operations to a complex function f: the Fourier transform F; the replacement of the Fourier-transform modulus by a function A obtained from experiments: AF[f]/|F[f]|; the inverse Fourier transform F^{-1} ; and the transformation P of f reflecting a priori information on the object. The result of these transformations is a new function

$$f' = P\left[F^{-1}\left[A\frac{F[f]}{|F[f]|}\right]\right]$$
(1)

defined in the same domain as f. If f' and f coincide with the specified accuracy, the goal is achieved. A priori information (P) can be the field modulus measured directly behind the object or in some other plane, the object shape, i.e., the zero field in the object plane outside some region, the nonnegativity of the field, etc.

We explicate the foregoing with the example of an error reduction algorithm used to determine the image of a transparent planar object. We assume that the object is illuminated by coherent radiation propagating through a round hole located directly in front of its surface. In this case, a priori information P is the zero field in the object plane outside the hole. We also assume that the field intensity distribution I is measured in a parallel plane at some distance from the object (in the detector plane). We also assume for simplicity that the field calculation in this plane reduces to the Fourier transform of the field in the object plane, and hence the modulus of the Fourier transform of the object field, A, is trivially calculated from \sqrt{I} . With these assumptions, the algorithm starts with the specification of an arbitrary complex function f within the hole including the object. At the first step, we perform the Fourier transformation F[f]corresponding to the calculation of the field propagation from the object to the detector. At the second step, the transformation AF[f]/|F[f]| is performed, which means that the modulus F[f] at each point of the inverse Fourier space is replaced with the corresponding value of A. The phase F[f] remains unchanged. At the third step, the inverse Fourier transformation $F^{-1}[AF[f]/|F[f]|]$ is performed, which corresponds to the inverse calculation of the field from the detector to the object. Finally, the last step is the transformation $P[F^{-1}(AF[f]/|F[f]|)]$, which means that the values of the complex function $F^{-1}[AF[f]/|F[f]|]$ outside the hole are assumed to be zero.

These four operations represent one iteration of the algorithm that gives a refined field distribution f' defined in the object plane. These iterations are performed during the algorithm operation until f' = f within a specified accuracy. As shown in [25], this algorithm always converges and is equivalent to the conjugate gradient method for seeking the field energy minimum outside the hole (a positive quantity, the so-called error), and is therefore called the error reduction algorithm. Unfortunately, such an algorithm almost always converges to a local minimum, preventing the reduction of the inherent error to zero. However, other, more sophisticated algorithms exist that can ensure the approach to the zero error. In the case of success, the function f found is treated as the problem solution. This function allows finding the amplitude and phase modulation by the object, which can be called the object image.

The scheme of a typical lensless diffraction-limited microscope contains a coherent radiation source illuminating an object that transmits radiation and is located in the plane \tilde{S} (Fig. 1). Unlike conventional microscopes, this scheme does not contain optical elements between the object and the detector. A laser illuminating the object on the left and computer processing of the data obtained from the detector are not shown in the diagram.

We show the role of the Fourier transform with the example of *a wave packet* that describes the propagation of coherent radiation from the object \tilde{S} to the detector S (Fig. 1) and is an exact solution of the wave equation for the field in the half-space in terms of spatial harmonics in the object



Figure 1. \hat{S} is the object plane, S is the detector plane, z is the distance between them, (L_x, δ_x) and (L_y, δ_y) are the detector domain size and pixel in directions x and y, respectively, and $(\tilde{L}_x, \tilde{\delta}_x)$ and $(\tilde{L}_y, \tilde{\delta}_y)$ are the object domain size and pixel in directions x and y, respectively.

plane:

$$\psi(\mathbf{r}) = \iint_{-\infty}^{\infty} \varphi_0(\mathbf{p}) d^2 \mathbf{p} \exp\left\{i\mathbf{p}\mathbf{\rho} + i\sqrt{k^2 - p^2}z\right\},$$

$$\mathbf{r} = (x, y, z) = (\mathbf{\rho}, z), \qquad (2)$$

where $k = 2\pi/\lambda$ is the wave vector and $\varphi_0(\mathbf{p})$ is the Fourier transform of the field distribution $\psi(\mathbf{p}, z = 0)$ in the object plane. Expression (2) was derived in [26–29], where its relation to other forms of the diffraction integral was considered. Here, we discuss only questions concerning simulations of the propagation of nonparaxial beams in problems of image phase retrieval and of ptychography. Wave packet (2), which is an exact solution of the wave equation, is free from limitations of the paraxial approximation and can be used to study effects near the optical-resolution limit associated with the wavelength.

We assume that a detector measuring the radiation intensity is located in the far-field zone, i.e., in the Fourier plane, where

$$z \gg \frac{a^2}{\lambda}, \quad \rho = z \tan \theta,$$
 (3)

a is the object size, and the aperture angle θ is fixed. Considering (2) under condition (3) and using the stationary phase method [30], we can readily obtain the field on the detector as

$$\psi(\mathbf{p}, z) = \frac{2\pi k}{\mathrm{i}z} \frac{z^2}{z^2 + \rho^2} \exp\left(\mathrm{i}k\sqrt{z^2 + \rho^2}\right) \\ \times \varphi_0\left(\mathbf{p} = k\frac{\mathbf{p}}{\sqrt{z^2 + \rho^2}}\right). \tag{4}$$

Expression (4), obtained from an exact solution of the Helmholtz wave equation, shows that the far-field spatial distribution is expressed in terms of the Fourier transform on the object. As the coordinate ρ on the detector changes from 0 to ∞ , the corresponding spatial harmonic p changes from 0 to k. In other words, the far-field zone does not contain harmonics with p > k. This expresses the general fact that the limit spatial resolution in the far-field zone is determined by the wavelength $\lambda = 2\pi/k$, irrespective of the optical system type.

In practice, because of difficulties in the numerical realization of algorithms based on exact expression (4), a simpler relation is used in phase-retrieval problems (see review [23]), the Fresnel integral,

$$\psi(\mathbf{\rho}, z) = \frac{k \exp\left(\mathrm{i}kz\right)}{2\pi \mathrm{i}z} \iint_{-\infty}^{\infty} \psi(\mathbf{\rho}', 0) \,\mathrm{d}^{2}\mathbf{\rho}' \exp\left(\mathrm{i}k \frac{(\mathbf{\rho} - \mathbf{\rho}')^{2}}{2z}\right),\tag{5}$$

which follows from exact expression (2) for small angles θ and in far-field zone (3) takes the form

$$\psi(\mathbf{p}, z) = \frac{2\pi k}{iz} \exp\left[ikz\left(1 + \frac{\rho^2}{2z^2}\right)\right] \varphi_0\left(\mathbf{p} = k\frac{\rho}{z}\right).$$
(6)

We compare expression (6) with the exact far-field field (4). A comparison of the arguments of the Fourier transform of the initial distribution $\varphi_0(\mathbf{p})$ shows that the Fourier integral in the far-field zone, Eqn (6), can be used to calculate the field on the detector only for $\rho \leq z$, i.e., for $\theta \leq 45^{\circ}$ and the aperture

 $NA = \sin \theta \le 0.71$. Otherwise, the harmonics of the initial distribution p > k would be involved, which, according to (4), are not contained in the far-field zone at all. By adding the requirement of the coincidence of phase factors in (4) and (6), we obtain the condition of the applicability of the Fresnel integral in the far-field zone (3) (see [31]):

$$\tan \theta \leqslant 1, \quad \frac{a^2}{\lambda} \ll z \leqslant \frac{4}{\pi} \frac{\lambda}{\tan^4 \theta} \,. \tag{7}$$

This problem is considered theoretically in more detail in review [32].

Thus, the wave packet approximation, like the paraxial approximation, leads in the far-field zone to the abovementioned relation between the field intensity on the detector and the modulus of the Fourier transform of the field in the object plane. Below, we call the wave packet approximation—in fact, the application of integral (2)—the wave packet method.³

3. Phase retrieval discretization in the wave packet method

We consider the setup of a lensless microscope in Fig. 1 assuming that the wavelength λ , the detector sizes L_x and L_y , the pixel sizes δ_x and δ_y , and the distance z between the object and the detector are specified in experiment. We then find the object sizes \tilde{L}_x and \tilde{L}_y (i.e., the field of view) and the spatial resolution $\tilde{\delta}_x$ and $\tilde{\delta}_y$. Different formulations of the problem are also possible, for example, if the size and scale of the resolved details of the object are specified and the distance z is used to optimize other parameters of the system presented in Fig. 1.

For the phase retrieval algorithms to work, the modulus of the Fourier transform should be known, which amounts to assuming the existence of a uniform rectangular frequency mesh on which the modulus is defined. Correspondingly, a rectangular uniform mesh on which the object field is defined is reciprocal to the frequency domain in which the Fouriertransform modulus is defined. We define the spatial and frequency domains as the sets

$$x|y \in \left\{\frac{-L_{x|y}}{2} + j\delta_{x|y}\right\}, \quad j = 0, \dots, N_{x|y},$$

$$p_{x|y} \in \left\{-\frac{\pi}{\delta_{x|y}} + j\frac{2\pi}{L_{x|y}}\right\}, \quad j = 0, \dots, N_{x|y},$$

$$\delta_{x|y} = \frac{L_{x|y}}{N_{x|y}},$$
(8)

where x|y means the direction along x or along y, $p_{x|y}$ is the spatial frequency, $L_{x|y}$, $N_{x|y}$, and $\delta_{x|y}$ are respectively the domain size, the number of pixels, and their size, determined by the physical properties of the detector, which are assumed to be known. It is assumed that the coordinate origin is located at the center (see Fig. 1).

The choice of the corresponding domain in the object plane is not unique. Neither its size nor the size of its pixels are known beforehand. We define the domain structure similarly to (8), specifying it by the spatial resolution $\delta_{x|y}$, the field of view $\tilde{L}_{x|y}$, and the number $\tilde{N}_{x|y}$ of pixels on the object.

³ Another term is the plane-wave angular spectrum method (https:// en.wikipedia.org.wiki/Angular_spectrum_method).

Hereafter, all the object parameters are indicated by a tilde. It follows from the results in Section 2 that the frequency domain of the object can be found from the known detector domain with the help of relation (4). Namely, the spatial frequency \tilde{p} of the object is related to the coordinate ρ on the detector as

$$\tilde{\mathbf{p}} = k \frac{\mathbf{\rho}}{\sqrt{z^2 + \rho^2}} \,. \tag{9}$$

Therefore, the maximum value is

$$\max\left(\tilde{p}_{x|y}\right) = k \frac{L_{x|y}/2}{\sqrt{z^2 + (L_{x|y}/2)^2}} \,. \tag{10}$$

On the other hand, as in (8), this value is $\pi/\delta_{x|y}$. This gives the optimum (largest) pixel size on the object

$$\tilde{\delta}_{x|y} = \frac{\pi}{\max\left(\tilde{p}_{x|y}\right)} = \frac{\lambda}{L_{x|y}} z \sqrt{1 + \left(\frac{L_{x|y}}{2z}\right)^2}.$$
(11)

The value $\delta_{x|y}$ is the maximum possible discretization step allowing the reconstruction of the object. In this case, the aperture of radiation emitted by an object part of size $\delta_{x|y}$ is completely overlapped by the detector. If the characteristic size of the object details is smaller than (11), the object reconstruction is impossible (if the object is continuous) and it is necessary to choose a shorter distance z. In the paraxial approximation $L_{x|y} \ll z$, expression (11) becomes

$$\tilde{\delta}_{x|y} = \frac{\lambda}{L_{x|y}} \, z \, .$$

Thus, the object pixel size for the wave packet method in Eqn (11) exceeds that in the paraxial approximation, and for large apertures $L_{x|y}/2z \ge 1$ tends to $\lambda/2$.

As in the derivation of (11), according (8), the mesh size in the object frequency domain is $2\pi/\tilde{L}_{x|y}$. At the same time, it follows from (9) that pixels of the object frequency domain are mapped into a pillow-shape mesh of image pixels on the detector. Because the gradient of the function $\tilde{p}(\rho)$ is

$$\frac{d\tilde{p}}{d\rho} = \frac{kz^2}{\left(z^2 + \rho^2\right)^{3/2}},$$
(12)

the density $\sigma_{x|y}(\rho)$ of these images per detector pixel in the direction x|y is given by

$$\sigma_{x|y}(\rho) = \frac{\mathrm{d}\tilde{p}_{x|y}/(2\pi/\tilde{L}_{x|y})}{\mathrm{d}(x|y)/\delta_{x|y}} = \frac{kz^2}{(z^2 + \rho^2)^{3/2}} \frac{\tilde{L}_{x|y}\delta_{x|y}}{2\pi}$$
$$= \frac{z^2}{(z^2 + \rho^2)^{3/2}} \frac{\tilde{L}_{x|y}\delta_{x|y}}{\lambda}.$$
(13)

Obviously, the equality $\sigma_{x|y}(\rho) = 1$ for all ρ would be ideal. This would correspond to the exact matching of two domains when their information capacities (the numbers of pixels) coincide. However, the density $\sigma_{x|y}(\rho)$ monotonically decreases with increasing ρ , having a maximum at the detector center $\rho = 0$:

$$\sigma_{x|y}(0) = \frac{\tilde{L}_{x|y}\delta_{x|y}}{z\lambda} \,. \tag{14}$$

If the size $\tilde{L}_{x|y}$ is chosen such that $\sigma_{x|y}(0)$ is less than unity, then the density of pixel images of the object on the detector is insufficient, which means a lower informative capacity of the object domain than that of the detector domain. Therefore, the inequality $\sigma_{x|y}(0) \ge 1$ should hold, which gives the minimal possible size of the object domain

$$\tilde{L}_{x|y} = \frac{\lambda}{\delta_{x|y}} z \,. \tag{15}$$

It is shown in Section 4 that such a choice of the object domain size provides good numerical accuracy, and we use it in what follows.

Expressions (11) and (15) determine the spatial resolution and the field of view in the object plane. For convenience, we present all the key formulas in the Table. With the notation $NA_{x|y} = \sin \theta_{x|y}$ and $\tan \theta_{x|y} = L_{x|y}/(2z)$, we obtain a classical formula for the diffraction-limited resolution from (11):

$$\tilde{\delta}_{x|y} = \frac{\lambda}{2\mathrm{NA}_{x|y}} \,. \tag{16}$$

Because we are considering the far-field zone, condition (3) should be satisfied with $a = \tilde{L}_{x|y}$. Taking (15) into account, we find from (3) that

$$z \leqslant \frac{\delta_{x|y}^2}{\lambda} \,. \tag{17}$$

We assume for simplicity that the detector has a square shape: $\delta_x = \delta_y = \delta$, $L_x = L_y = L$, and $N_x = N_y = N$. It then follows from (11) and (15) that the domain in the object plane is also square: $\tilde{\delta}_x = \tilde{\delta}_y = \tilde{\delta}$, $\tilde{L}_x = \tilde{L}_y = \tilde{L}$, $\tilde{N}_x = \tilde{N}_y = \tilde{N}$. According to (15) and (17), the distance

$$z_{\rm m} = \frac{\delta^2}{\lambda} \tag{18}$$

exists for a square domain at which the field of view \tilde{L} reaches the maximum

$$\tilde{L}_{\rm m} = \delta \tag{19}$$

for a given detector. Thus, when the detector is located in the far-field zone, the field of view on the object is determined by

Table. Formulas for determining the spatial resolution $\delta_{x|y}$ and the field of view $\tilde{L}_{x|y}$ depending on $(L_{x|y}, \delta_{x|y})$ for the detector, the wavelength λ , and the distance z. Square brackets denote the integral part.

$\tilde{\delta}_{x y}$ (11)	$\tilde{L}_{x y}$ (15)	$ ilde{N}_{x y} = [ilde{L}_{x y}/ ilde{\delta}_{x y}]$
$\frac{\lambda}{L_{x y}} z \sqrt{1 + \left(\frac{L_{x y}}{2z}\right)^2}$	$rac{\lambda}{\delta_{x y}}$ Z	$\left[\left(\frac{L_{x y}}{\delta_{x y}}\right)/\sqrt{1+\left(\frac{L_{x y}}{2z}\right)^2}\right] = \left[N_{x y}/\sqrt{1+\left(\frac{L_{x y}}{2z}\right)^2}\right]$

the detector pixel size. Substituting z_m in (11), we obtain the pixel size and the number of pixels on the object for the maximal field of view:

N L Popov, I A Artyukov, A V Vinogradov, V V Protopopov

$$\tilde{\delta}_{\rm m} = \frac{\delta^2}{L} \sqrt{1 + \left(\frac{\lambda L}{2\delta^2}\right)^2},\tag{20}$$

$$\tilde{N}_{\rm m} = \frac{\delta}{\tilde{\delta}_{\rm m}} = \left[\frac{L}{\sqrt{\delta^2 + (\lambda L/2)^2 (1/\delta^2)}}\right].$$
(21)

It follows from (21) that the function $\tilde{N}_{\rm m}(\delta)$ has a maximum at $\delta = \delta_{\rm M} = \sqrt{\lambda L/2}$, with

$$\tilde{N}_{\rm M} = \left[\sqrt{\frac{L}{\lambda}}\right], \ \tilde{\delta}_{\rm M} = \frac{\lambda}{\sqrt{2}}, \ \tilde{L}_{\rm M} = \sqrt{\frac{\lambda L}{2}}, \ z_{\rm M} = \frac{L}{2}.$$
 (22)

Thus, by setting δ equal to $\delta_{\rm M}$, we maximize \tilde{N} and hence the amount of information obtained. In this case, it is obvious that the aperture is

$$NA_M = \sin \theta_M = \frac{1}{\sqrt{2}} \approx 0.71, \text{ i.e. } \theta_M = 45^\circ.$$
 (23)

This aperture provides the optimum combination of high resolution with a large field of view (Fig. 2). We note that the maximum possible number of pixels on the object is independent of the pixel size on the detector, but is determined only by the detector size and the wavelength.

In practice, the required detector pixel size can be achieved by combining adjacent physical pixels into one virtual pixel. Such an operation is called data binning—the summation of signals from adjacent cells of the detector into one pixel. For example, if four adjacent square cells are combined into one virtual pixel, then, according to (19), the field of view increases twofold. In this case, the distance z_m (see (18)) increases fourfold, while the resolution $\delta_{\rm m}$ deteriorates, according to (20). Other than the increase in the field of view, binning allows expanding the dynamic range of the number of incident photons per pixel and increasing the signal-to-noise ratio.

It follows from (22) that the maximum possible number $\tilde{N}_{\rm M}$ of pixels on the object increases with decreasing λ . For example, for $\lambda = 0.5 \ \mu m$ and $L = 1 \ cm$, we have $N_{\rm M} = 141$, while for $\lambda = 10$ nm, $N_{\rm M} = 10^3$. Thus, to obtain resolution in the visible range close to the diffraction limit, we should reduce the number of pixels on the object and the related field of view $\tilde{L}_{\rm M} = \tilde{N}_{\rm M} \delta_{\rm M}$. However, ptychography (see Section 5) provides an increase in the field of view due to a large number of shots without deteriorating the spatial resolution.

The argument concerning the optimal field of view presented above is applied to a classical lensless object reconstruction from one shot. In the case of ptychography, the argument regarding the digital domain and the optimal size of an object, Eqn (19) applies to each individual exposure (scan).

Because relation (9) between $\tilde{\mathbf{p}}$ and \mathbf{p} is nonlinear, a uniform lattice in the ρ space is mapped into a nonuniform lattice in the $\tilde{\mathbf{p}}$ space. However, to use the discrete Fourier transform, the latter should also be uniform. Therefore, the interpolation of the function $|\varphi_0(\tilde{\mathbf{p}}(\mathbf{p}))|$ is necessary. Fortunately, this function is quite smooth in practice, and therefore its interpolation poses no problem. However, the

Figure 2. Dependence of $\tilde{N}/\tilde{N}_{\rm M}$ on the dimensionless quantity $2\delta/\lambda$. The maximum is reached at $\delta = \delta_{\rm M} = \lambda/\sqrt{2}$.

calculation of the complex amplitude $\psi(\mathbf{p}, z)$ in the direct problem by expression (4), as in the numerical experiment considered in Section 4, requires a full interpolation, which can be performed by expanding in a Taylor series in the vicinity of a point nearest to $\rho(\tilde{\mathbf{p}})$. In this case, derivatives can be calculated using the fast Fourier transform as $-\partial_{mn}\varphi_0(\tilde{\mathbf{p}}(\mathbf{\rho})) = F^{(-1)}[(i\tilde{p}_x)^m(i\tilde{p}_y)^n F]\varphi_0(\tilde{\mathbf{p}}(\mathbf{\rho})).$ For example, to achieve the result presented in Section 4, it was necessary to calculate 20 additional discrete Fourier transforms (derivatives through the fifth order). In the general case, obviously, the interpolation reduces the accuracy and increases the calculation time. Therefore, whenever possible, it is natural to use the paraxial approximation, in which interpolation is absent (because $\tilde{\mathbf{p}} = k\mathbf{p}/z$ depends linearly on ρ). Obviously, this somewhat impairs the spatial resolution δ .

4. Results of calculations. Comparison of the wave packet and paraxial approximation methods

To estimate the accuracy of calculations by wave-packet formulas (4) and in paraxial approximation (6) under the condition that the digital domain in the object plane is calculated by expressions (11) and (15), we made a comparison with a point-like source function:

$$\psi_{s}(\mathbf{r}) = \frac{1}{|\mathbf{r}|} \exp\left(ik|\mathbf{r}|\right),$$

$$\varphi_{0s}(\mathbf{p}) = \frac{i}{2\pi\sqrt{k^{2} - p^{2}}}.$$
(24)

It can be readily shown that the substitution of (24) in (4)leads to the identity, i.e., the wave-packet approximation in the far-field zone gives the exact result.

Numerical calculations were performed for the wavelength $\lambda = 10$ nm, and domains in the object and detector planes were chosen as in Section 3. The detector pixel size was $\delta = 13 \,\mu\text{m}$, and the distance to the detector and its size were chosen in accordance with (18) and (22), i.e., $z = z_m =$ 1.69 cm, $L = 2z_m = 3.38$ cm. In this case, the resolution is $\delta_{\rm M} = 7$ nm.

Numerical calculations have shown that the wave-packet method gives the relative error $|(\psi(\mathbf{r}) - \psi_s(\mathbf{r}))/\psi_s(\mathbf{r})| \sim 10^{-16}$ in the angular interval $\tan \theta \in [-1, 1]$, except for a small number of angles (artifacts) at which the error is of the order

Physics-Uspekhi 63 (8)



of 10^{-7} . The accuracy of the paraxial approximation rapidly decreases with increasing θ , such that the relative error exceeds 0.013 already for $\tan \theta > 0.01$. According to the estimate of the applicability of paraxial approximation (7), the angle θ must be quite small: $\tan \theta < \sqrt[4]{4\lambda/(\pi z)} = 0.029$. The calculation showed that the relative error of the paraxial approximation for $\tan \theta = 0.029$ was 0.88.

Thus, the numerical experiment with point-like source function (24) shows that the error of the paraxial approximation exceeds 50% already for the angle $\theta = 1.66^{\circ}(\tan \theta = 0.029)$, whereas the numerical calculation by the wave-packet method (4) actually gives the exact result. In the region $0 < \theta < 26^{\circ}$, the error is within 10^{-16} , which corresponds to the machine epsilon. In the region $26^{\circ} < \theta < 45^{\circ}$, the error remains at the same level for most of the points and is about 10^{-7} for the rest of the points.

5. Ptychography. Results of calculations

Ptychography is a method of computer imaging of overlapping field intensities (scans) recorded upon small lateral displacements of an object in the plane S (Fig. 1). During the object scan, the illumination beam and the detector position are assumed invariable. The use of several overlapping scans instead of one scan (as in reconstructing the phase) allows dropping a priori conditions related to the object properties. The principles of ptychography and the term 'ptychography' itself were proposed in [21, 22]. However, this approach remained poorly known for a long time because of the absence of an efficient algorithm for its realization. The first such algorithm was proposed by Rodenburg only in 2004 [33, 34] after his long-term studies beginning in the early 1990s.⁴ This algorithm was called the Ptychography Iterative Engine (PIE). However, the algorithm had a disadvantage: the field of a source illuminating the object was assumed to be known beforehand. After four years, in 2008, this disadvantage was eliminated in another iterative algorithm based on the difference-map approach [35, 36]. Somewhat later, in 2009, an improved PIE algorithm was proposed-the extended Ptychography Iterative Engine (ePIE) [37], which we use in this paper to illustrate ptychography.

The ePIE algorithm is used to solve the system of J equations

$$A_{j}(\tilde{\mathbf{\rho}}) = \left| FFT \left[P\left(\tilde{\mathbf{\rho}} - \tilde{\mathbf{\rho}}_{j} \right) O(\tilde{\mathbf{\rho}}) \right] \right|, \ j = 1, \dots, J.$$
(25)

Here, $P(\tilde{\mathbf{\rho}} - \tilde{\mathbf{\rho}}_i)$ is the amplitude of the illumination beam on the object surface ('probe function'), $O(\tilde{\mathbf{p}})$ is the required transmission (or reflection) function of the object, $P(\tilde{\rho} - \tilde{\rho}_i)O(\tilde{\rho})$ is the amplitude of the transmitted (reflected) wave on the object surface, $A_i(\tilde{\rho})$ is the modulus of the Fourier transform of the wave propagating from the object on its surface (cf. the beginning of Section 2), J is the number of scans used in the calculation, j is the number of a scan, and $\{\tilde{\mathbf{\rho}}_i\}_{i=1}^J$ is the known set of displacements for which $P(\tilde{\mathbf{\rho}} - \tilde{\mathbf{\rho}}_i)$ and $P(\tilde{\mathbf{p}} - \tilde{\mathbf{p}}_{i-1})$ overlap by 60–70%. The last is necessary for the unique solution of system of equations (25). The use of the probe function is a distinct feature of ptychography. The probe function can be known a priori, found simultaneously with the object function $O(\tilde{\mathbf{p}})$, or found independently by standard phase retrieval methods. It is important that it be unchanged during measurements [37]. In more detail, the ePIE algorithm involves the iterative simultaneous calculation of the object function $O(\tilde{\rho})$ and the probe function $P(\tilde{\rho})$:

$$O_{j+1}(\tilde{\boldsymbol{\rho}}) = O_j(\tilde{\boldsymbol{\rho}}) + \alpha \frac{P_j^*(\tilde{\boldsymbol{\rho}} - \tilde{\boldsymbol{\rho}}_j)}{\left|P_j(\tilde{\boldsymbol{\rho}} - \tilde{\boldsymbol{\rho}}_j)\right|_{\max}^2} \left(f_j'(\tilde{\boldsymbol{\rho}}) - f_j(\tilde{\boldsymbol{\rho}})\right), \quad (26)$$

$$P_{j+1}(\tilde{\mathbf{p}}) = P_j(\tilde{\mathbf{p}}) + \beta \frac{O_j^*(\tilde{\mathbf{p}} + \tilde{\mathbf{p}}_j)}{\left|O_j(\tilde{\mathbf{p}} + \tilde{\mathbf{p}}_j)\right|_{\max}^2} \left(f_j'(\tilde{\mathbf{p}}) - f_j(\tilde{\mathbf{p}})\right), \qquad (27)$$

$$O_1(\tilde{\mathbf{\rho}}) = O_0(\tilde{\mathbf{\rho}}), \qquad (28)$$

$$P_1(\tilde{\mathbf{\rho}}) = P_0(\tilde{\mathbf{\rho}}), \qquad (29)$$

$$f_j(\tilde{\mathbf{\rho}}) = O_j(\tilde{\mathbf{\rho}}) P_j(\tilde{\mathbf{\rho}} - \tilde{\mathbf{\rho}}_j), \qquad (30)$$

$$f_{j}'(\tilde{\mathbf{\rho}}) = F^{-1} \left[A_{j}(\tilde{\mathbf{\rho}}) \frac{F\left[f_{j}(\tilde{\mathbf{\rho}})\right]}{\left|F\left[f_{j}(\tilde{\mathbf{\rho}})\right]\right|} \right],$$
(31)

$$A_{j+J}(\tilde{\mathbf{\rho}}) = A_j(\tilde{\mathbf{\rho}}), \ \tilde{\mathbf{\rho}}_{j+J} = \tilde{\mathbf{\rho}}_j, \tag{32}$$

where α and β are dimensionless coefficients of the order of unity, $O_0(\tilde{\mathbf{p}})$ is the initial object, $P_0(\tilde{\mathbf{p}})$ is the initial probe function, and *j* is the iteration number. Scans are cyclically repeated with the period *J*. The process stops when $O_j(\tilde{\mathbf{p}})$ ceases to change. We note that the only quantity in (26)–(32) directly related to the detector signal is the modulus of the field amplitude $A_j(\tilde{\mathbf{p}})$. All the other quantities are determined in calculations.

We assumed in our numerical experiment that $\alpha = 1$ and $\beta = 1$. The real probe function $P(\tilde{\mathbf{p}})$ was chosen equal to the field of a point-like source located at a distance of 10 cm in a round hole with diameter \tilde{L} . Displacements were chosen over a spiral:

$$\tilde{\rho}_{j} = \frac{j}{128} \, 18.2 \left(1 + \text{rand} \left(-0.05, \ 0.05 \right) \right) \quad [\mu\text{m}] \,,$$

$$\tilde{\varphi}_{j} = \frac{2\pi}{16} \, j \left(1 + \text{rand} \left(-0.05, \ 0.05 \right) \right), \ j = 0, \dots, 127 \,,$$
(33)

where rand (-0.05, 0.05) is a random number from -0.05 to 0.05. Altogether, there are J = 128 overlapping regions of the object, each 13 µm in diameter. The total field of view is close to a circle 49.4 µm in diameter (Fig. 3).

The digital domain \tilde{L} , $\tilde{\delta}$ considered in Section 3 should be used only to calculate (31); it is there that the function $A_j(\tilde{\rho})$ is defined. We call this domain the Fourier domain. Because the field of view is 3.8 times larger than the Fourier domain $(13 \times 13 \ \mu\text{m}^2)$, we chose the digital domain of a larger size, $7 \times \tilde{L}$, with the same pixel $\tilde{\delta}$. This size is just sufficient to accommodate all the displacements of the Fourier domain in expressions (26)–(30), (33). This domain was used for the functions $O_j(\tilde{\rho})$, $P_j(\tilde{\rho})$, $f_j(\tilde{\rho})$, and $f'_j(\tilde{\rho})$, while the initial function $O_0(\tilde{\rho})$ was a Gaussian with $\sigma = 2 \ \mu\text{m}$.

The wavelength was 10 nm. A square-shape detector with square pixels with $\delta = 13 \ \mu m$ was used. The object was a fractal template (Fig. 4) $50 \times 50 \ \mu m^2$ in size. The height of numbers in this template is equal to the number value multiplied by 50 nm, i.e., the height of the largest number '100' is 5 $\ \mu m$ and the height of the smallest number '2' is 100 nm. The white and black colors respectively correspond to 1 and 0.

We make preliminary estimates of the geometry and some details of the numerical experiment. According to (18) and

⁴ The history of this question is described in more detail in review [24].



Figure 3. Total field of view consisting of the combination of 128 disks 13 μ m in diameter arranged over a spiral. The boundaries of the disks are shown. The field of view close in shape to a disk 49.5 μ m in diameter is inserted into the inner square with a side of 50 μ m. The large square with a side of 91 μ m shows the general domain used for calculations.



Figure 4. Object in the form of a fractal template $50 \times 50 \ \mu\text{m}^2$ in size. The height of numbers is the 'number value' $\times 50$ nm, i.e., the height of the largest number '100' is 5000 nm, and the height of the smallest number '2' is 100 nm. The white and black colors respectively correspond to 1 and 0.

(23), to obtain the maximum amount of information, we should choose $z_{\rm m} = 1.69$ cm, L = 3.38 cm, and N = 2600. Then, according to (22), the resolution and the field of view are $\delta_{\rm M} = 7$ nm and $\tilde{L}_{\rm M} = 13$ µm. However, a physical pixel of the detector measures a discrete quantity, the number of photons, which cannot be less than zero. This means that the number of photons in one shot should be large enough to be sufficient for pixels remote from the axis. The initial numerical calculation of the field on the detector by



expression (4) for the object (see Fig. 4), the probe function $P(\tilde{\mathbf{p}})$, and the geometric parameters z_m and L presented above with N = 2600 showed that the number of photons per pixel near the axis should be ~ 10⁸. This greatly exceeds the saturation threshold of a standard silicon pixel (~ 10³) at this wavelength. To take this feature into account, we changed the numerical experiment geometry by introducing the 16 × 16 binning of detector pixels with a simultaneous 16-fold increase in the distance. According to (15) and (16), the size of \tilde{L} remained the same (13 µm), while the resolution deteriorated to $\tilde{\delta} = 80$ nm. In order to 'see' '2' on the

deteriorated to $\delta = 80$ nm. In order to see 2 on the reconstructed object, we increased the detector size and improved the resolution to $\tilde{\delta} = 34$ nm. Thus, our parameters were z = 27 cm, L = 8 cm, N = 385, $\delta = 208 \ \mu\text{m}$, $\tilde{L} = 13 \ \mu\text{m}$, and $\tilde{\delta} = 34$ nm. The photon distribution obtained on the detector is shown in Fig. 5. The maximum value is 1.75×10^6 , which corresponds to ~ 7000 photons per physical pixel.

At the first stage, we calculated the $O(\tilde{\mathbf{p}})P(\tilde{\mathbf{p}}-\tilde{\mathbf{p}}_j)$ field distributions on the detector in accordance with expression (4) for all the 128 { $\tilde{\mathbf{p}}_j$ } displacements (33). The probe function $P(\tilde{\mathbf{p}})$ was normalized such that its modulus squared was equal to the number of photons incident on the unit area, and the field intensity integral over a pixel on the detector gave the number of incident photons. After that, the rounding to integers and noise contamination by the Poisson distribution were performed. Then the reverse calculation of the modulus of the object Fourier transform $A_j(\tilde{\mathbf{p}})$ given the photon distribution on the detector was performed by expression (4) (see Fig. 5). In this case, the simplest zero-order interpolation was used.

At the second stage, ptychography algorithm (26)–(32) was used. After the 47,360th iteration (370 cycles over 128 scans), the reconstructed object and the probe function ceased to change and acquired the form shown in Fig. 6, where the object is shown on the left and the probe function is shown on the right. Looking at the probe function, we can see that part of the lower edge and part of the right edge are slightly 'bitten off'. This is related to the invariance of expressions (26)-(32) under the displacement of the object coordinate system. Therefore, a small displacement of the reconstructed object and the probe function with respect to their real positions is quite expected. Despite this, we can assume that the object reconstruction problem has been successfully solved in this case. The central parts of the reconstructed (Fig. 7a) and initial (Fig. 7b) images $10 \times 10 \ \mu\text{m}^2$ in size are compared in Fig. 7. The smallest number '2' on the reconstructed image can be distinguished



Figure 6. Reconstructed images of the object (left) and the illumination function (right) after the 47,360th iteration (370 cycles over 128 scans). The object diameter is 50 μ m, the illumination function diameter is 13 μ m.



Figure 7. Comparison of the central parts of (a) the reconstructed image and (b) the original.

but cannot be read. This corresponds to the claimed resolution of 34 nm at which 3×3 pixels represent the number '2', which is insufficient for recognition but is sufficient for detection.

The time spent on the image reconstruction itself from the first iteration to the 47,360th iteration was 43 min or 0.05 s per iteration. A 14-core CPU Intel(R) Core(TM) i9-7940X CPU, 3.10 GHz, 128 Gb computer was used.

6. Conclusions

Ptychography has been transformed in recent years into an all-wavelength microscopy method providing accurate phase information without using high-resolution optical elements and a priori information on an object. Ptychography is applied in the X-ray, VUV, and visible wavelength ranges. Thus, we can say that the idea of lensless optics [1] proposed about half a century ago has been realized in practice. Commercial 'ptychoscopes' for cytology have appeared [17–20]. The experimental system of a lensless microscope in the simplest case includes four elements: a coherent radiation source, an object stage, scanning the object perpendicular to the optical axis, and a detector with a computer for scan processing (Fig. 8).

In this paper, we have given a brief introduction to the method of ptycho-imaging. As the main diffraction integral relating the field distributions on the object and on the detector, we use a wave packet in the far-field zone. This



Figure 8. Ptychoscope setup.

preserves the possibility of imaging with the diffractionlimited resolution $\tilde{\delta} = \lambda/2$. At the same time, calculation algorithms still involve the Fourier transform of fields on the object and detector. We have obtained theoretically substantiated expressions determining the size, the discretization step of a domain on the object, and the spatial resolution. The distance between the object and the detector, the size of the detector, and the pixel size are determined by experimental conditions and are assumed to be known. For a square-shape detector, criteria are formulated for choosing the optimal distance and the size of detector pixels (binning).

The accuracy of expressions for choosing the object domain has been demonstrated for a point-like source. Numerical ptychographic experiments on microscopy have demonstrated the possibility of imaging with a $50 \times 50 \ \mu\text{m}^2$ field of view and a resolution of 34 nm at the wavelength 10 nm. The reconstruction time for one object was 43 min of personal computer operation and can be reduced by 2 to 3 orders of magnitude with the help of appropriate software and computers.

The further development of ptychography requires the expansion of its applications and the experimentally based systematic analysis of the accuracy and stability of the image retrieval algorithms. This is necessary in order to realize the possibilities of ptychography as a unique microscopic method with the diffraction-limited resolution determined by the wavelength.

Acknowledgments

The authors are grateful to the participants of the seminars led V S Lebedev and I V Smetanin for the fruitful discussions.

This study was supported by grants 19-02-00394, 18-08-01066, 18-29-17039, and 17-08-01286; basic funding within the framework of theme 0023-0002-2018, and the Program of Scientific Studies of the Presidium of RAS "Actual Problems of Photonics, Probing of Inhomogeneous Media and Materials" (PP-RAS no. 7).

References

- 1. Gerchberg R W, Saxton W O Optik 35 237 (1972)
- Brady G R, Guizar-Sicairos M, Fienup J R Opt. Express 17 624 (2009)
- 3. Katkovnik V et al. Optica 4 786 (2017)

- 4. McDermott S, Maiden A Opt. Express 26 25471 (2018)
- 5. Sandberg R L et al. Proc. Natl. Acad. Sci. USA 105 24 (2008)
- Miao J, Sandberg R L, Song C IEEE J. Select. Topics Quantum Electron. 18 399 (2012)
- 7. Zürch M et al. Sci. Rep. 4 7356 (2015)
- 8. Chapman H N, Nugent K A Nat. Photon. 4 833 (2010)
- 9. Robisch A-L et al. Opt. Lett. **41** 5519 (2016)
- 10. Kuznetsova T I Sov. Phys. Usp. **31** 364 (1988); Usp. Fiz. Nauk **154** 677 (1988)
- 11. Bakut P A et al. Sov. Phys. Dokl. **31** 710 (1986); Dokl. Akad. Nauk SSSR **290** 89 (1986)
- Ustinov N D, Matveev I N, Protopopov V V Metody Obrabotki Opticheskikh Polei v Lazernoi Lokatsii (Methods for Optical Field Processing and Laser Location) (Moscow: Nauka, 1983)
- Bel'dyugin I M, Zubarev I G, Mikhailov S I Quantum Electron. 31 539 (2001); Kvantovaya Elektron. 31 5329 (2001)
- Kotlyar V V, Malov A N Sov. J. Quantum Electron. 19 696 (1989); Kvantovaya Elektron. 16 5 1072 (1989)
- Goncharskii A V, Goncharskii A A Komp'yuternaya Optika. Komp'yuternaya Golografiya (Computer Optics. Computer Holography) (Moscow: Izd. MGU, 2004)
- Karasik V E, Stasenko K V, Tolstoguzov V L Vestn. Mos. Gos. Tech. Univ. im. N.E. Baumana (Special Issue) 214 (2011) Special Issue: Modern Problems of Optical Engineering
- 17. Suman R et al. Sci. Rep. 6 22032 (2016)
- 18. Marrison J et al. *Sci. Rep.* **3** 2369 (2013)
- 19. Rath N et al. Cancer Res. 78 3321 (2018)
- 20. Phase Focus Limited, https://www.phasefocus.com/
- 21. Hoppe W Acta Cryst. A 25 495 (1969)
- 22. Hegerl R, Hoppe W Ber. Bunsen. Phys. Chem. 74 1148 (1970)
- 23. Shechtman Y et al. IEEE Signal Process. Mag. 32 (3) 87 (2015)
- 24. Pfeiffer F Nat. Photon. **12** 9 (2018)
- 25. Fienup J R Appl. Opt. 21 2758 (1982)
- 26. Lalor É J. Opt. Soc. Am. 58 1235 (1968)
- 27. Goodman J W Introduction to Fourier Optics 2nd ed. (New York: McGraw-Hill, 1996)
- 28. Onural L J. Opt. Soc. Am. A 28 290 (2011)
- Kim M K Digital Holographic Microscopy: Principles, Techniques, and Applications (Springer Series in Optical Sciences, Vol. 162) (New York: Springer, 2011)
- Fedoryuk M V Metod Perevala (Method of Steepest Descent) (Moscow: Librokom, 2010)
- 31. Vinogradov A V et al. *Numerical Simulation of X-ray Diffraction Optics* (Moscow: A and B Publ. House, 1999)
- Kiselev A P Opt. Spectrosc. 102 603 (2007); Opt. Spektrosk. 102 4661 (2007)
- 33. Faulkner H M L, Rodenburg J M Phys. Rev. Lett. 93 023903 (2004)
- 34. Rodenburg J M, Faulkner H M L Appl. Phys. Lett. 85 4795 (2004)
- 35. Thibault P et al. *Science* **321** 379 (2008)
- 36. Thibault P et al. Ultramicroscopy 109 338 (2009)
- 37. Maiden A M, Rodenburg J M Ultramicroscopy 109 1256 (2009)