$\mu(xy)$ which describes phenomenologically the absorption in thin specimens. Then, instead of (2.5.5.5), the wave on the exit surface of a specimen can be written as

$$q(xy) = 1 - i\sigma\varphi(xy) - \mu(xy) \tag{2.5.5.12}$$

and in the back focal plane if $\Phi = \mathscr{F}\varphi$ and $M = \mathscr{F}\mu$

$$Q(uv) = \delta(uv) - i\sigma\Phi(uv) - M(uv). \tag{2.5.5.13}$$

Usually, μ is small, but it can, nevertheless, make a certain contribution to an image. In a sufficiently good linear approximation, it may be assumed that the real part $\cos\chi$ of the phase function (2.5.5.7*a*) affects $M(uv)$, while $\Phi(xy)$, as we know, is under the action of the imaginary part $\sin\chi$.

Thus, instead of (2.5.5.6), one can write

$$Q(\exp i\chi) = \delta(\mathbf{u}) - i\sigma\Phi(\mathbf{u})\sin\chi - M(\mathbf{u})\cos\chi, \tag{2.5.5.14}$$

and as the result, instead of (2.5.5.10),

$$\begin{aligned} I(xy) = 1 &+ 2\sigma\varphi(xy) * \mathscr{F}^{-1}(\sin\chi) * a(U) \\ &- 2\mu(xy) * \mathscr{F}^{-1}(\cos\chi) * a(U). \end{aligned} \tag{2.5.5.15}$$

The functions $\varphi(xy)$ and $\mu(xy)$ can be separated by object imaging using the through-focus series method. In this case, using the Fourier transformation, one passes from the intensity distribution (2.5.5.15) in real space to reciprocal space. Now, at two different defocus values Δf_1 and Δf_2 [(2.5.5.6), (2.5.5.7*a*,*b*)] the values $\Phi(\mathbf{u})$ and $M(\mathbf{u})$ can be found from the two linear equations (2.5.5.14). Using the inverse Fourier transformation, one can pass on again to real space which gives $\varphi(\mathbf{x})$ and $\mu(\mathbf{x})$ (Schiske, 1968). In practice, it is possible to use several through-focus series and to solve a set of equations by the least-squares method.

Another method for processing takes into account the simultaneous presence of noise $N(\mathbf{x})$ and transfer function zeros (Kirkland *et al.*, 1980). In this method the space frequencies corresponding to small values of the transfer function modulus are suppressed, while the regions where such a modulus is large are found to be reinforced.

2.5.5.4. Thick crystals

When the specimen thickness exceeds a certain critical value (~50–100 Å), the kinematic approximation does not hold true and the scattering is dynamic. This means that on the exit surface of a specimen the wave is not defined as yet by the projection of potential $\varphi(xy) = \int\varphi(\mathbf{r})\,\mathrm{d}z$ (2.5.5.3), but one has to take into account the interaction of the incident wave ψ_0 and of all the secondary waves arising in the whole volume of a specimen.

The dynamic scattering calculation can be made by various methods. One is the multislice (or phase-grating) method based on a recurrent application of formulae (2.5.5.3) for n thin layers Δz_i thick, and successive construction of the transmission functions q_i (2.5.5.4), phase functions $Q_i = \mathscr{F}q_i$, and propagation function $p_k = [k/2\pi i\Delta z]\exp[ik(x^2 + y^2)/2\Delta z]$ (Cowley & Moodie, 1957).

Another method – the scattering matrix method – is based on the solution of equations of the dynamic theory (Chapter 5.2). The emerging wave on the exit surface of a crystal is then found to diffract and experience the transfer function action [(2.5.5.6), (2.5.5.7*a*,*b*)].

The dynamic scattering in crystals may be interpreted using Bloch waves:

$$\Psi^j(\mathbf{r}) = \sum_H C_H^j \exp(-2\pi i\mathbf{k}_H^j \cdot \mathbf{r}). \tag{2.5.5.16}$$

It turns out that only a few (bound and valence Bloch waves) have strong excitation amplitudes. Depending on the thickness of a crystal, only one of these waves or their linear combinations (Kambe, 1982) emerges on the exit surface. An electron-microscopic image can be interpreted, at certain thicknesses, as an image of one of these waves [with a correction for the transfer function action (2.5.5.6), (2.5.5.7*a*,*b*)]; in this case, the identical images repeat with increasing thickness, while, at a certain thickness, the contrast reversal can be observed. Only the first Bloch wave which arises at small thickness, and also repeats with increasing thickness, corresponds to the projection of potential $\varphi(xy)$, *i.e.* the atom projection distribution in a thin crystal layer.

An image of other Bloch waves is defined by the function $\varphi(\mathbf{r})$, but their maxima or minima do not coincide, in the general case, with the atomic positions and cannot be interpreted as the projection of potential. It is difficult to reconstruct $\varphi(xy)$ from these images, especially when the crystal is not ideal and contains imperfections. In these cases one resorts to computer modelling of images at different thicknesses and defocus values, and to comparison with an experimentally observed pattern.

The imaging can be performed directly in an electron microscope not by a photo plate, but using fast-response detectors with digitized intensity output online. The computer contains the necessary algorithms for Fourier transformation, image calculation, transfer function computing, averaging, and correction for the observed and calculated data. This makes possible the interpretation of the pattern observed directly in experiment (Herrmann *et al.*, 1980).

2.5.5.5. Image enhancement

The real electron-microscope image is subdivided into two components:

$$J(xy) = I(xy) + N(xy). \tag{2.5.5.17}$$

The main of these, $I(xy)$, is a two-dimensional image of the 'ideal' object obtained in an electron microscope with instrumental functions inherent to it. However, in the process of object imaging and transfer of this information to the detector there are various sources of noise. In an electron microscope, these arise owing to emission-current and accelerating-voltage fluctuations, lens-supplying current (temporal fluctuations), or mechanical instabilities in a device, specimen or detector (spatial shifts). The two-dimensional detector (*e.g.* a photographic plate) has structural inhomogeneities affecting a response to the signal. In addition, the specimen is also unstable; during preparation or imaging it may change owing to chemical or some other transformations in its structure, thermal effects and so on. Biological specimens scatter electrons very weakly and their natural state is moist, while in the electron-microscope column they are under vacuum conditions. The methods of staining (negative or positive), *e.g.* of introducing into specimens substances containing heavy atoms, as well as the freeze-etching method, somewhat distort the structure of a specimen. Another source of structure perturbation is radiation damage, which can be eliminated at small radiation doses or by using the cryogenic technique. The structure of stained specimens is affected by stain graininess. We assume that all the deviations $\Delta I_k(xy)$ of a specimen image from the 'ideal' image $I_k(xy)$ are included in the noise term $N_k(xy)$. The substrate may also be inhomogeneous. All kinds of perturbations cannot be separated and they appear on an electron microscope image as the full noise content $N(xy)$.

The image enhancement involves maximum noise suppression $N(xy)$ and hence the most accurate separation of a useful signal $I(xy)$ from the real image $J(xy)$ (2.5.5.1). At the signal/noise ratio $I/N \simeq 1$ such a separation appears to be rather complicated. But in some cases the real image reflects the structure sufficiently well, *e.g.* during the atomic structure imaging of some crystals ($I/N > 10$). In other cases, especially of biological specimen imaging, the noise N distorts substantially the image, (I/N) ~ 5–10. Here one should use the methods of enhancement. This problem is usually solved by the methods of statistical processing of sets of images J_k ($k = 1, \ldots, n$). If one assumes that the informative signal $I_k(xy)$ is always the same, then the noise error $N(xy)$ may be reduced.

The image enhancement methods are subdivided into two classes:

(a) image averaging in real space xy;

(b) Fourier analysis and filtration in reciprocal space.

These methods can be used separately or in combination. The enhancement can be applied to both the original and the restored images; there are also methods of simultaneous restoration and enhancement.

The image can be enhanced by analogue (mainly optical and photographic) methods or by computational methods for processing digitized functions in real and reciprocal space.

The cases where the image has translational symmetry, rotational symmetry, and where the image is asymmetric will be considered.

Periodic images. An image of the crystal structure with atomic or molecular resolution may be brought to self-alignment by a shift by a and b periods in a structure projection. This can be performed photographically by printing the shifted image on the same photographic paper or, *vice versa*, by shifting the paper (McLachlan, 1958).

The Fourier filtration method for a periodic image I_p with noise N is based on the fact that in Fourier space the components $\mathscr{F}I_p$ and $\mathscr{F}N$ are separated. Let us carry out the Fourier transformation of the periodic signal I_p with the periods a, b and noise N:

$$\begin{aligned}\mathscr{F}J &= \mathscr{F}[I_p(xy) + N(xy)] \\ &= \textstyle\int I_p(xy) \exp[2\pi i(hx + ky)]\, \mathrm{d}x\, \mathrm{d}y + \mathscr{F}N \\ &= \textstyle\sum \Phi_{hk} \delta(\mathbf{u} - \mathbf{u}_{hk}) + \mathscr{F}N; \qquad (2.5.5.18) \\ \mathbf{u}_{hk} &= h\mathbf{a}^* + k\mathbf{b}^*.\end{aligned}$$

The left part of (2.5.5.18) represents the Fourier coefficients Φ_{hk} distributed discretely with periods a^* and b^* in the plane $\mathbf{u}(uv)$. This is the two-dimensional reciprocal lattice. The right-hand side of (2.5.5.18) is the Fourier transform $\mathscr{F}N$ distributed continuously in the plane. Thus these parts are separated. Let us 'cut out' from distribution (2.5.5.18) only Φ_{hk} values using the 'window' function $w(uv)$. The window should match each of the real peaks Φ_{hk} which, owing to the finite dimensions of the initial periodic image, are not points, as this is written in an idealized form in (2.5.5.18) with the aid of δ functions. In reality, the 'windows' may be squares of about $a^*/10$, $b^*/10$ in size, or a circle. Performing the Fourier transformation of product (2.5.5.18) without $\mathscr{F}N$, and set of windows $w(\mathbf{u}) = w(uv) * \sum_{h,k} \delta(\mathbf{u} - h\mathbf{a}^* - k\mathbf{b}^*)$, we obtain:

$$\begin{aligned}J(xy) &= \mathscr{F}^{-1}\{w(\mathbf{u}) \sum_{h,k} \Phi_{h,k} \delta(\mathbf{u} - \mathbf{u}_{h,k})\} \\ &= W(xy) * I_p(\mathbf{x}), \qquad (2.5.5.19)\end{aligned}$$

the periodic component without the background, $W(xy) = \mathscr{F}^{-1}w(\mathbf{u})$. The zero coefficient Φ_{00} in (2.5.5.19) should be decreased, since it is due, in part, to the noise. When the window w is sufficiently small, I_p in (2.5.5.19) represents the periodic distribution $\langle I \rangle$ (average over all the unit cells of the projection) included in I_p (2.5.5.18). Nevertheless, some error from noise in an image does exist, since with Φ_{hk} we also introduced into the inverse Fourier transformation the background transform values $\mathscr{F}^{-1}N_{hk}$ which are within the 'windows'.

This approach is realized by an analogue method [optical diffraction and filtering of electron micrographs in a laser beam (Klug & Berger, 1964)] and can also be carried out by computing.

As an example, Fig. 2.5.5.2(b) shows an electron micrograph of the periodic structure of a two-dimensional protein crystal, while Fig. 2.5.5.2(c) represents optical diffraction from this layer. In order to dissect the aperiodic component $\mathscr{F}N$ in a diffraction plane, according to the scheme in Fig. 2.5.5.2(a), one places a mask with windows covering reciprocal-lattice points. After such a filtration, only the I_p component makes a contribution during the image formation by means of a lens, while the component $\mathscr{F}N$ diffracted by the background is delayed. As a result, an optical pattern of the periodic structure is obtained (Fig. 2.5.5.2d).

Optical diffractometry also assists in determining the parameters of a two-dimensional lattice and its symmetry.

Using the same method, one can separate the superimposed images of two-dimensional structures with different periodicity and in different orientation, the images of the 'near' and 'far' sides of tubular periodic structures with monomolecular walls (Klug & DeRosier, 1966; Kiselev *et al.*, 1971), and so on.

Computer filtering involves measuring the image optical density J_{obs}, digitization, and Fourier transformation (Crowther & Amos, 1971). The sampling distance usually corresponds to one-third of the image resolution. When periodic weak phase objects are investigated, the transformation (2.5.5.18) yields the Fourier coefficients. If necessary, we can immediately make corrections in them using the microscope transfer function according to (2.5.5.6), (2.5.5.7a,b) and (2.5.5.11a), and thereby obtain the true kinematic amplitudes Φ_{hk}. The inverse transformation (2.5.5.16) gives a projection of the structure (Unwin & Henderson, 1975; Henderson & Unwin, 1975).

Sometimes, an observed image $J(\mathbf{x})$ is 'noised' by the $N(\mathbf{x})$ to a great extent. Then, one may combine data on real and reciprocal space to construct a sufficiently accurate image. In this case, the electron-diffraction pattern is measured and structure-factor moduli from diffraction reflection intensities $I_{hk,\,obs}$ are obtained:

$$|\Phi_{hk,\,obs}| \sim \sqrt{I_{hk,\,obs}}. \qquad (2.5.5.20)$$

At the same time, the structure factors

$$\Phi_{hk,\,calc} = |\Phi_{hk,\,calc}| \exp(i\alpha_{hk,\,calc}) \qquad (2.5.5.21)$$

are calculated from the processed structure projection image by means of the Fourier transformation. However, owing to poor image quality we take from these data only the values of phases α_{hk} since they are less sensitive to scattering density distortions than the moduli, and construct the Fourier synthesis

$$\begin{aligned}I(xy) = \sum_{hk} &|\Phi_{hk,\,obs}| \exp(i\alpha_{hk,\,calc}) \\ &\times \exp[2\pi i(hx + ky)]. \qquad (2.5.5.22)\end{aligned}$$

Here the possibilities of combining various methods open up, *e.g.* for obtaining the structure-factor moduli from X-ray

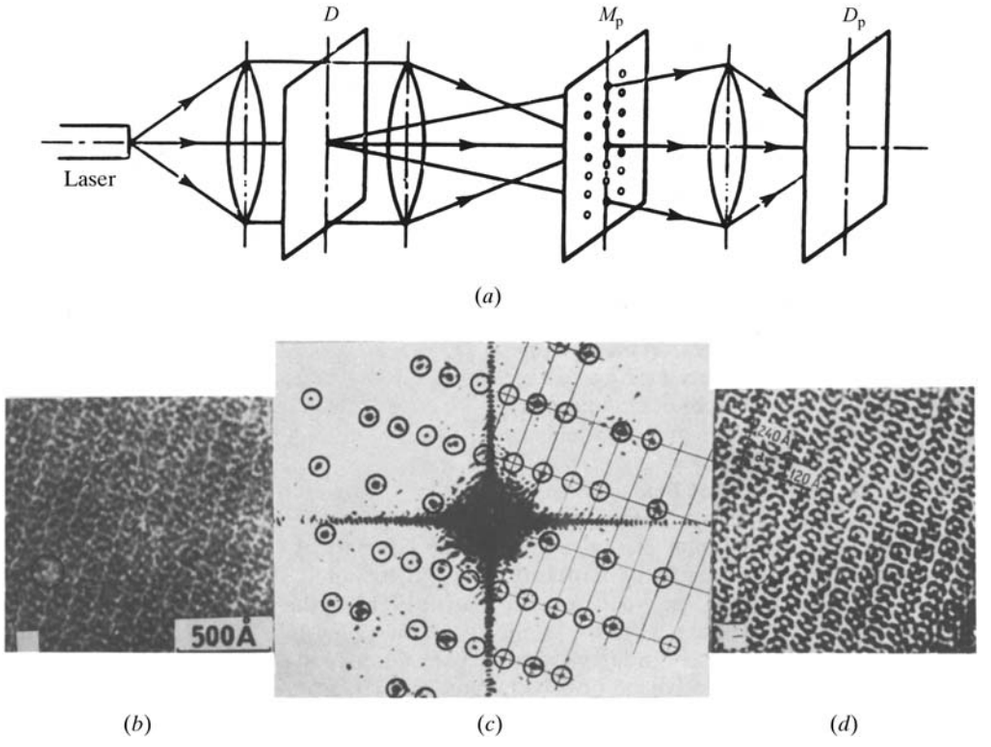

Fig. 2.5.5.2. (*a*) Diagram of an optical diffractometer. *D* is the object (an electron micrograph), M_p is the diffraction plane and a mask that transmits only Φ_{hk}, D_p is the plane of the (filtered) image; (*b*) an electron micrograph of a crystalline layer of the protein phosphorylase *b*; (*c*) its optical diffraction pattern (the circles correspond to the windows in the mask that transmits only the Φ_{hk} diffracted beams from the periodic component of the image); (*d*) the filtered image. Parts (*b*)–(*d*) are based on the article by Kiselev *et al.* (1971).

diffraction, and phases from electron microscopy, and so on (Gurskaya *et al.*, 1971).

Images with point symmetry. If a projection of an object (and consequently, the object itself) has a rotational *N*-fold axis of symmetry, the structure coincides with itself on rotation through the angle $2\pi/N$. If the image is rotated through arbitrary angles and is aligned photographically with the initial image, then the best density coincidence will take place at a rotation through $\alpha = (k2\pi/N)$ $(k = 1, \ldots, N)$ which defines *N*. The pattern averaging over all the rotations will give the enhanced structure image with an $(N)^{1/2}$ times reduced background (Markham *et al.*, 1963).

Rotational filtering can be performed on the basis of the Fourier expansion of an image in polar coordinates over the angles (Crowther & Amos, 1971).

$$I(r, \psi) = \sum_{n=-\infty}^{+\infty} g_n(r) \exp(in\varphi). \qquad (2.5.5.23)$$

The integral over the radius from azimuthal components g_n gives their power

$$p_n \sim \int_0^a |g_n|^2 r \, \mathrm{d}r, \qquad (2.5.5.24)$$

where *a* is the maximum radius of the particle. A set p_n forms a spectrum, the least common multiple *N* of strong peaks defining the *N*-fold symmetry. The two-dimensional reconstructed image of a particle with rotational symmetry is defined by the synthesis (2.5.5.24) with $n = 0, N, 2N, 3N$.

Asymmetric images. In this case, a set of images is processed by computational or analogue methods. The initial selection of images involves the fulfilment of the maximum similarity condition.

The averaging of *n* images in real space gives

$$I_{\text{enh}} = (1/n) \sum_{k=1}^{n} J_k(xy) = \langle I_k \rangle(xy) + (1/n) \sum N_k(xy). \qquad (2.5.5.25)$$

The signal/noise ratio on an average image is $(n)^{1/2}$ times enhanced.

The degree of similarity and accuracy of superposition of two images with an account both of translational and angular shifts is estimated by a cross-correlation function[4] of two selected images J_1 and J_2 (Frank, 1975, 1980).

$$\begin{aligned} k(\mathbf{x}') &= J_1 * J_2 = \int J_1(x) J_2(x + x') \, \mathrm{d}x \\ &= k_{I_1 I_2} + k_{I_1 N_2} + k_{I_2 N_1} + k_{N_1 N_2}. \end{aligned} \qquad (2.5.5.26)$$

The value $k(\mathbf{x}')$ is the measure of image similarity, the x' coordinate of the maximum indicates the shift of the images relative to each other. The first term of the resultant expression (2.5.5.26) is the cross-correlation function of noise-corrected images being compared, the second and third terms are approximately equal to zero, since the noise does not correlate with the signal; the last term is the autocorrelation function of the noise (Cramér, 1954; Frank, 1975, 1980).

The calculation of a correlation function is performed by means of Fourier transformation on the basis of the convolution theorem, since the Fourier transformation of the product of the Fourier transform of function J_1 and the conjugated Fourier transform function J_2 gives the cross-correlation function of the initial functions:

[4] At $I_j = I_k$ this is the autocorrelation function, an analogue of the Patterson function used in crystallography.

$$k = \mathscr{F}^{-1}[\mathscr{F}J_1 \cdot \mathscr{F}^* J_2]. \tag{2.5.5.27}$$

The probability density of samples for images has the form

$$p(J_1 J_2 \ldots J_n) = \frac{1}{(\sigma\sqrt{2\pi})^n} \times \exp\left[\frac{-1}{2\sigma^2}\sum^{n}\int [J_k(\mathbf{x}+\mathbf{x}_k) - J(\mathbf{x})]^2 \, \mathrm{d}x\right]. \tag{2.5.5.28}$$

Here J is the tentative image (as such, a certain 'best' image can first be selected, while at the repeated cycle an average image is obtained), $J_k(\mathbf{x})$ is the image investigated, σ is the standard deviation of the normal distribution of noises and x_k the relative shift of the image. This function is called a likelihood function; it has maxima relative to the parameters $J(x)$, x_k, σ. The average image and dispersion are

$$J(\mathbf{x}) = (1/n)\sum^{n}[J_k(\mathbf{x}-\mathbf{x}_k)],$$
$$\sigma^2 = (1/n)\sum^{n}[J_k(\mathbf{x}-\mathbf{x}_k) - J(\mathbf{x})]^2. \tag{2.5.5.29}$$

This method is called the maximum-likelihood method (Cramér, 1954; Kosykh *et al.*, 1983).

It is convenient to carry out the image alignment, in turn, with respect to translational and angular coordinates. If we start with an angular alignment we first use autocorrelation functions or power spectra, which have the maximum and the symmetry centre at the origin of the coordinates. The angular correlation maximum

$$f(\theta') = \int f_k(\theta - \theta') f_e(\theta) \, \mathrm{d}\theta \tag{2.5.5.30}$$

gives the mutual angle of rotation of two images.

Then we carry out the translational alignment of rotationally aligned images using the translational correlation function (2.5.5.26) (Langer *et al.*, 1970).

In the iteration alignment method, the images are first translationally aligned and then an angular shift is determined in image space in polar coordinates with the centre at the point of the best translational alignment. After the angular alignment the whole procedure may be repeated (Steinkilberg & Schramm, 1980).

The average image obtained may have false high-frequency components. They can be excluded by multiplying its Fourier components by some function and suppressing high-space frequencies, for instance by an 'artificial temperature factor' $\exp\{-B|\mathbf{u}|^2\}$.

For a set of similar images the Fourier filtration method can also be used (Ottensmeyer *et al.*, 1977). To do this, one should prepare from these images an artificial 'two-dimensional crystal', *i.e.* place them in the same orientation at the points of the two-dimensional lattice with periods a, b.

$$J = \sum_{k=1}^{n} J_k(\mathbf{x} - \mathbf{t}_p); \ \mathbf{t} = p_1\mathbf{a} + p_2\mathbf{b}. \tag{2.5.5.31}$$

The processing is then performed according to (2.5.5.18), (2.5.5.19); as a result one obtains $\langle I(xy)\rangle$ with reduced background. Some translational and angular errors in the arrangement of the images at the artificial lattice points act as an artificial temperature factor. The method can be realized by computing or by optical diffraction.

2.5.6. Three-dimensional reconstruction[5]

By B. K. Vainshtein and P. A. Penczek

2.5.6.1. The object and its projection

In electron microscopy (EM) and single-particle reconstruction, three-dimensional (3D) reconstruction methods are used for studying biological structures; that is, symmetric or asymmetric associations of biomacromolecules (muscles, spherical and rod-like viruses, bacteriophages, individual proteins and ribosomes) (Frank, 2006). The electron microscope is used to obtain parallel-beam two-dimensional (2D) projections $\varphi_2(\mathbf{x}, \boldsymbol{\tau})$ (ray transform) of frozen hydrated 3D macromolecules $\varphi_3(\mathbf{r})$ suspended in random orientations (Fig. 2.5.6.1). The function $\varphi_2(\mathbf{x}_\tau)$ is the 2D projection of the 3D molecular electron distribution $\varphi_3(\mathbf{r})$. One can also consider one-dimensional (1D) projections $\varphi_1(s, \boldsymbol{\tau})$ of multidimensional distributions; the set of these projections is called a Radon transform. For 2D distributions, ray and Radon transforms differ only in the notation. For ($d > 2$)-dimensional distributions the two are different: in a ray transform the integrals are calculated over straight lines and yield ($d - 1$)-dimensional projections, while in Radon transforms the integrals are calculated over ($d - 1$)-dimensional hyperplanes and yield 1D projections. In electron microscopy, Radon transforms are not directly measurable, but can be formed computationally and used in intermediate steps of the 3D reconstruction or in alignment procedures (Radermacher, 1994).

Within the linear weak-phase-object approximation of the image-formation process in the microscope [see equation (2.5.2.42) in Section 2.5.2 of this chapter], 2D projections represent line integrals of the potential of the particle under examination convoluted with the point-spread function of the microscope, s, so, using (2.5.2.43),

$$I(xy) = 1 + 2\sigma s(xy) * \varphi(xy) = 1 + 2\sigma s(xy) * \int \varphi_3(\mathbf{r}) \, \mathrm{d}z. \tag{2.5.6.1}$$

Since

$$\varphi_2(\mathbf{x}_\tau) = \int \varphi_3(\mathbf{r}) \, \mathrm{d}\tau, \quad \boldsymbol{\tau} \perp \mathbf{x}, \tag{2.5.6.2}$$

we have

$$I(xy) = 1 + 2\sigma s(\mathbf{x}_\tau) * \varphi_2(\mathbf{x}_\tau), \quad \boldsymbol{\tau} \perp \mathbf{x}. \tag{2.5.6.3}$$

If we omit constant terms, we obtain

$$I(xy) = s(\mathbf{x}_\tau) * \varphi_2(\mathbf{x}_\tau) = s(\mathbf{x}_\tau) * \int \varphi_3(\mathbf{r}) \, \mathrm{d}\tau, \quad \boldsymbol{\tau} \perp \mathbf{x}. \tag{2.5.6.4}$$

In this section, we will assume that all images were collected using the same defocus setting, so the point-spread function s is constant and does not depend on the projection direction $\boldsymbol{\tau}$. Thus, we will concern ourselves with the inversion of the projection problem

$$\varphi_2(\mathbf{x}_\tau) = \int \varphi_3(\mathbf{r}) \, \mathrm{d}\tau, \quad \boldsymbol{\tau} \perp \mathbf{x}. \tag{2.5.6.5}$$

[5] The original version of Section 2.5.6, written by the late B. K. Vainshtein, is here updated and expanded by P. A. Penczek.