VOLTERRA TRANSFER MODELS FOR ELECTRON WAVE RECONSTRUCTION

Adriaan van den Bos

Department of Applied Physics,
Delft University of Technology,
P.O. Box 5046, 2600 GA Delft, The Netherlands,
email vdbos@tn.tudelft.nl

ABSTRACT

For the transmission electron microscope, the transfer of the complex specimen exit wave to observed intensities is modelled as a Volterra system function. This system function includes the effect of microscope fluctuations and is used in a parameter estimation based reconstruction of the exit wave. The numerical optimization employed in the parameter estimation procedure is carried out directly with respect to both the real and the complex valued parameters.

1. INTRODUCTION

In optics, Volterra kernels or rather their Fourier transforms, the Volterra system functions, are called transmission cross coefficients and are used to describe combinations of spatially static nonlinear and dynamic linear operations on complex waves [1]. The spatial coordinate is usually two-dimensional.

In this paper, the system function of the transmission electron microscope (TEM) is discussed. The purpose is to show how the system function can be used for reconstruction of the specimen exit wave. This is the complex valued spatial wave at the exit interface of the specimen investigated. It is used in the assessing of material structure. The relation of the observed image in the TEM and the exit wave to be reconstructed is as follows. The magnifying system of the TEM is modelled as a two-dimensional linear transfer function. The effect of this linear transfer is called blurring. The detection of the magnified wave is a modulus square operation. The result of this operation, in which the phase of the wave is lost, is the detected TEM image. Thus the detected image consists of intensities observed as an electron count per pixel. The counts are supposed to be Poisson distributed.

The waves considered in this paper are spatially periodic since the specimen is supposed to be a homogeneous crystal. In the chosen reconstruction method, the periodicity is used as follows. Since the wave is periodic, it is defined by its Fourier coefficients. Therefore, the reconstruction may be reformulated as estimating the parameters, that is, the Fourier coefficients of the wave from the error corrupted image observations. From the estimated Fourier coefficients the estimate of the wave is obtained by Fourier synthesis. As a result of the microscope properties, the spatial bandwidths are strictly limited and, therefore, the total number of Fourier coefficients is bounded. For various reasons, in this frequency domain parameter estimation (FDPE) method [2] and in alternative methods proposed in the literature [3]-[6], the wave is reconstructed from a series of images. Each of the images is measured at a different defocus. Since the transfer function depends on the defocus and this dependence is known, each of the images corresponds to a different known transfer function. The series of images is called defocus series. The linear transfer and the modulus square operation combined represent a homogeneous second order Volterra system function which is known but different for each defocus. This known system function operates on the wave Fourier coefficient spectrum. Thus for each defocus the observations are a known quadratic function of the unknown, complex Fourier coefficients of the wave. These functions are fitted simultaneously for all defoci. If this fitting is carried out in the least squares sense, the least squares criterion is a quartic polynomial in the complex Fourier coefficients sought and consists of as many partial sums as there are defoci employed. The main advantage of the FDPE approach is that the number of parameters is small and independent of the number of pixels processed. In the methods proposed in the literature, each of the many pixels processed is an unknown. In addition, in the FDPE approach use can be made of the statistical knowledge about the pixels to define maximum likelihood estimators which are asymptotically most precise. Thus the FDPE method solves, in principle, the periodic wave reconstruction from error corrupted observations.

The FDPE method sketched assumes the defoci employed to be constants. In practice, they fluctuate and, as a result, so do the transfer functions. The FDPE method has in common with the most important reconstruction method currently in use [3] that these defocus fluctuations are not included in the transfer model. In [3] this is corrected for by weighted addition of reconstruction results according to the probability of the defocus values. Here, a different approach is followed. Use is made of the fact that the average effect of the defocus fluctuations can be included in closed form in the Volterra system function [7]. The including does not complicate the estimation procedure. For example, the least squares criterion remains a numerically well-behaved quartic polynomial. This is important since in FDPE the least squares method is always used to either produce the final estimates or the initial values for the numerical maximization of the likelihood function.

The parameters to be estimated in the FDPE method are the Fourier coefficients and possibly the spatial periods of the complex exit wave. These quantities are all complex valued, except the constant term of the wave and the periods. Therefore, the likelihood function and the least squares criterion are both real functions to be optimized with respect to a mixture of real and complex variables. For that purpose, a special Newton method has been developed. It is an extension of the purely complex Newton method proposed in [8]. The advantages of this mixed real complex optimization over purely real optimization with respect to the real variables and the separate real and imaginary parts of the complex variables are a drastic simplification of the resulting expressions for the Hessian matrix and the gradient employed and the resulting simplicity of the codes.

2. NONLINEAR TRANSFER AND IMAGING IN THE TEM

Suppose that the spatially periodic, complex valued exit wave of the specimen in a TEM is described by

$$f(x,y) = \sum_{k,l} \gamma_{kl}\, e^{j\,2\pi\,(\nu_k\, x + \nu_l\, y)} \tag{1}$$

Here, x and y are the spatial coordinates, $j = \sqrt{-1}$ and ν_k and ν_l are equal to k/ξ and l/η where ξ and η are the periods in the x and the y direction, respectively, k and l are the harmonic numbers and the γ_{kl} are the Fourier coefficients which are all complex except the constant term γ_{00}. The spectrum of the γ_{kl} is not Hermitian since $f(x,y)$ is complex valued. The exit wave is magnified by a linear system with a two-dimensional linear transfer function $H(k,l,D)$ which blurs the wave. The variable D represents the defocus of the TEM. The result after magnification is

$$g(x,y,D) = \sum_{k,l} H(k,l,D)\, \gamma_{kl}\, e^{j\,2\pi\,(\nu_k\, x + \nu_l\, y)} \tag{2}$$

The image is the intensity of this wave as a function of x and y. It is described by

$$I(x,y,D) = |g(x,y,D)|^2 \tag{3}$$

Then from (2) and (3) it follows that it is equal to

$$\begin{aligned} I(x,y,D) \;=\; & \sum_{k_1,l_1} \sum_{k_2,l_2} \gamma_{k_1 l_1}\, \gamma^*_{k_2 l_2} \\ & \times H(k_1,l_1,D)\, H^*(k_2,l_2,D) \\ & \times e^{j\,2\pi\{(\nu_{k_1}-\nu_{k_2})x + (\nu_{l_1}-\nu_{l_2})y\}} \end{aligned} \tag{4}$$

which represents a homogeneous second-order Volterra description [9] of the nonlinear transfer in the TEM. Leaving out the frequency independent magnification, a general description of the system function $H(k,l,D)$ of a TEM is

$$H(k,l,D) = e^{j\,(\alpha_{kl} + \beta_{kl} D)} \tag{5}$$

with $\alpha_{kl} = \frac{\pi}{2} C_s \lambda^3 \left(\nu_k^2 + \nu_l^2\right)^2$ and $\beta_{kl} = \pi\lambda\left(\nu_k^2 + \nu_l^2\right)$ where C_s is the spherical aberration constant and λ is the electron wave length [7].

The observations $w(p,q)$ of the image are made in a number of discrete points (x_p, y_q), $p = 1,...,P$ and $q = 1,...,Q$. These observations are electron counts which are assumed to be Poisson distributed stochastic variables. Their expectations $\mathcal{E}\left[w(p,q)\right]$ are equal to the exact image $I(x_p, y_q, D)$ described by (4). Furthermore, the observations are usually made such that they may be considered to be statistically independent. The defocus series of images is made at the defoci D_m, $m = 1,...,M$. The problem solved by the FDPE method may be summarized as estimating the Fourier coefficients γ_{kl} from observations $w(p,q,m)$ with $p = 1,...,P$, $q = 1,...,Q$ and $m = 1,...,M$ where the index m has been added to define the defocus value employed.

3. DEFOCUS SPREAD AND SYSTEM FUNCTION

In the previous section the defocus value was considered a constant. However, even if the fluctuations of the voltages and currents in the TEM would be negligible, there would still be defocus fluctuations, called defocus spread, resulting from energy spread of the electrons. Typically, these fluctuations have a bandwidth of tens of MHz. They may be incorporated in the considerations of the previous section by everywhere substituting $D+d$ for D where D is now the expectation of the defocus and d is the zero mean fluctuation superimposed on it. Generally, it is assumed that the d are normally distributed while their standard deviation σ is quite accurately known. By Eq. (5)

$$H(k,l,D+d) = H(k,l,D)\, e^{j\,\beta_{kl} d} \tag{6}$$

This equation shows that the system function has become a stochastic variable itself. As a result the intensity described by Eq. (4) becomes $I(x,y,D+d)$, which is a stochastic variable. Equation (4) shows that the computation of the expectation $\mathcal{E}_d\left[I(x,y,D+d)\right]$ requires the computation of

$$\mathcal{E}_d\left[H(k_1,l_1,D+d)\; H^*(k_2,l_2,D+d)\right] =$$

$$H(k_1,l_1,D)\; H^*(k_2,l_2,D)\; \mathcal{E}_d\left[e^{j\,\Delta\beta_{12} d}\right] \tag{7}$$

with $\Delta\beta_{12} = \beta_{k_1 l_1} - \beta_{k_2 l_2} = \pi\lambda\left(\nu_{k_1}^2 + \nu_{l_1}^2 - \nu_{k_2}^2 - \nu_{l_2}^2\right)$ where use has been made of Eq. (6). The expression $\mathcal{E}_d\left[e^{j\,\Delta\beta_{12}\, d}\right]$ is recognized as the value of the characteristic function of d at the point $\Delta\beta_{12}$. Therefore, since d is normally distributed, it is equal to $e^{-\frac{1}{2}(\Delta\beta_{12}\sigma)^2}$ [10]. Combining this result with Eqs (4) and (7) yields

$$I(x,y,D) = \mathcal{E}_d\left[I(x,y,D+d)\right] =$$

$$= \sum_{k_1,l_1}\sum_{k_2,l_2} \gamma_{k_1 l_1}\gamma_{k_2 l_2}^* \times$$

$$\times H(k_1,l_1,D)\; H^*(k_2,l_2,D)$$

$$\times e^{j\,2\pi\left\{(\nu_{k_1}-\nu_{k_2})x+(\nu_{l_1}-\nu_{l_2})y\right\}}$$

$$\times e^{-\frac{1}{2}\pi^2\lambda^2\sigma^2\left(\nu_{k_1}^2+\nu_{l_1}^2-\nu_{k_2}^2-\nu_{l_2}^2\right)^2} \tag{8}$$

The last exponential factor in the summand of this expression results from the defocus spread. It is seen to increasingly limit the bandwidth as the standard deviation of the defocus variations increases. Let the quantity $I(x_p,y_q,D_m)$ represent the expectation of the image observations $w(p,q,m)$ where the index m has been added as a reference to the defocus D_m. Then an observation $w(p,q,m)$ is equal to the sum of $I(x_p,y_q,D_m)$ and a zero mean stochastic error. After replacing the γ_{kl} in Eq.(8) by corresponding variables c_{kl}, the model is obtained to be fitted with respect to the variables c_{kl} to the observations $w(p,q,m)$. This will be discussed in the next section.

4. MIXED REAL COMPLEX PARAMETER ESTIMATION

Equations (4) and (8) describe the expectations of the image with and without inclusion of defocus spread, respectively. Comparing these equations shows that the including of the defocus does not seriously complicate the model to be fitted. It is still a homogeneous second order Volterra model. If least squares model fitting is chosen, the criterion to be minimized is described by

$$\sum_m\sum_{p,q}\left\{w(p,q,m) - I(x_p,y_q,D_m)^2\right\} \tag{9}$$

This is a quartic multivariable polynomial in the Fourier coefficients c_{kl}. The coefficient c_{00} is real. All other coefficients are complex valued. Therefore, the least squares criterion described by Eq. (9) is a real function of one real variable and a number of complex variables. If, in addition, the periods ξ and η are unknown, they have to be estimated along with the Fourier coefficients. Then the number of real parameters increases to three. Let the vector $r = \begin{bmatrix} X & Y & c_{00}\end{bmatrix}^T$ or, if the periods are known, $r = c_{00}$ represent the real parameters in these cases, where X and Y are variables corresponding to the periods ξ and η, respectively. Furthermore, let vector c be the vector of all complex c_{kl} in both cases. Then the estimating problem could be reformulated as the estimation of r and all real and imaginary parts of the elements of c. The disadvantage of this approach is that it leads to complicated expressions. Here a simpler alternative is proposed: the direct fitting of the model described by Eq.(8) with respect to the mixture of real and complex parameters X, Y, c_{00} and the elements of c. This is a relatively straightforward extension of the purely complex optimization discussed in [8]. Define the mixed real complex vector of parameters t as

$$t = \begin{pmatrix} r^T & c^T & c^H \end{pmatrix}^T \tag{10}$$

where the superscripts T and H denote transposition and complex conjugate transposition, respectively. Notice that both the elements of c and their complex conjugates are present in t. Suppose that $f(t)$ is a real function of the elements of t such as the least squares criterion. Then the vector $\partial f(t)/\partial t$ is defined as the complex gradient vector of $f(t)$ [8]. The matrix of second order derivatives of $f(t)$ is defined as $\partial^2 f(t)/\partial t^*\,\partial t^T$ [8] where the superscript $*$ denotes complex conjugation. Using this gradient and Hessian matrix, the mixed real complex Newton optimization step Δt_N becomes [8]

$$\Delta t_N = -\left(\frac{\partial^2 f}{\partial t^*\,\partial t^T}\right)^{-1}\left(\frac{\partial f}{\partial t^*}\right) \tag{11}$$

where, for simplicity, the argument of $f(t)$ has been omitted. In the same way, other gradient methods for numerical optimization can be transformed into mixed real complex ones. Currently, the software for wave reconstruction using real parameter estimation described in [2] is rewritten in this mixed real complex form.

5. CONCLUSIONS

It has been shown that the nonlinear transfer of the complex exit wave to the detected image in the transmission electron microscope can be described by a relatively simple, homogeneous second order Volterra system function. Using this system function, a straightforward method for reconstruction of the wave has been proposed based on parameter estimation. It has also been shown that defocus spread can be incorporated in the system function. Compared to existing reconstruction methods this is a simplification. A further, numerical, simplification has been achieved by estimating the complex valued parameters of the wave directly instead of by estimating separately their real and imaginary parts.

6. REFERENCES

[1] M. Born and E. Wolf, *Principles of Optics*, New York: Pergamon, 1980.

[2] A. van den Bos, "Complex electron wave reconstruction using parameter estimation," accepted for publication in *IEEE Transactions on Instrumentation and Measurement.*

[3] E.J. Kirkland, "Improved high resolution image processing of bright field electron micrographs," *Ultramicroscopy*, vol. 15, pp. 151-172, 1984.

[4] W. Coene, G. Janssen, M. Op de Beeck and D. van Dyck, "Phase retrieval through focus variation for ultra-resolution in field-emission transmission electron microscopy," *Physical Review Letters*, vol. 69, pp. 3743-3746, 1992.

[5] D. van Dyck and M. Op de Beeck, "A new approach to object wavefunction reconstruction in electron microscopy," *Optik*, vol. 93, pp. 103-107, 1993.

[6] M.A.O. Miedema, A. van den Bos, and A.H. Buist, "Experimental design of exit wave reconstruction from a transmission electron microscope defocus series," *IEEE Transactions on Instrumentation and Measurement*, vol. 43, pp.181-186, 1994.

[7] K. Ishizuka, "Contrast transfer of images in TEM," *Ultramicroscopy*, vol. 5, pp. 55-65, 1980.

[8] A. van den Bos, "Complex gradient and Hessian," *IEE Proceedings Vision, Image and Signal Processing*, vol. 141, pp. 380-382, 1994.

[9] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, New York: Wiley, 1980.

[10] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill, 1991.