Based on the Schrödinger equation, the multislice method is derived in a new way which is more appropriate for calculating coherent STEM images. It is shown that the effects of beam tilt can be included in two equivalent ways which are convenient for different purposes of application. It is shown that, if the coherent STEM probe is considered as a superposition of many tilted plane waves, one cannot use the traditional multislice formulae for each tilted beam, since this leads to the loss of phase information which can be important for coherent many-beam illumination. The new formulations lead to a faster procedure for coherent STEM image simulations. Advantages of the proposed procedure are discussed with respect to a few different illumination conditions.

1. Introduction

Among current practical simulations of high-energy electron diffraction and imaging, the STEM image simulation may well be the most time-consuming one [1].

In principle, there are two ways for calculating a STEM image with multislice methods [2]. One is that proposed by Spence [3] for calculating CBED patterns of dislocations. In this method, a periodic continuation of the crystal unit cell has to be made to include sufficient sampling points and the probe function is treated as a whole and serves as the first phase grating of the multislice formula [4]. This is up to now the most popular multislice procedure for calculating coherent STEM images and CBED patterns. Another way for calculating STEM images and CBED patterns with the multislice method is that first calculations are separately made for each direction within the incident cone and then intensities (for incoherent illumination) or amplitudes with the relative phases (for coherent cases) are superimposed. Currently, this procedure is mainly used for incoherent cases [5]. This procedure would also be profitable for coherent cases, if amplitudes with the relative phases could be obtained for each incident beam. However, in the traditional multislice formula for treating beam tilt [6, 7], some of the relative phases seem to be lost [8].

In the present work more complete multislice expressions for treating beam tilt are derived from the basic Schrödinger equation, and based thereupon a faster procedure (as compared with the conventional one) is proposed for coherent STEM image simulations. Advantages

---

1) Groenenborgerlaan 171, B-2020 Antwerp, Belgium.
2) Postdoc Researcher for the National Fund for Scientific Research (Belgium).
3) Research Assistant for the National Fund for Scientific Research (Belgium).
of the new procedure are discussed with respect to a few different illumination conditions. Some results comparing the practical times between the traditional procedure and the new procedure are also given.

2. Formulations of the Multislice Method for Coherent Many-Beam Illumination

The multislice method, which was originally derived from optical principles by Cowley and Moodie [9], has alternatively been derived from the quantum mechanical principles (e.g., [10 to 14]). For our purpose, formulations of the multislice method are derived by following the real-space approach as has been proposed by Van Dyck [15] but keeping the complete phase information in the derivation, so as to obtain a new expression of the multislice method, which is convenient for treating coherent many-beam illumination.

In this paper, we use the following symbols to represent wave functions and their Fourier transforms:

\[ \Psi(R) \leftrightarrow \Psi(K) \]

for the complete wave function and

\[ \Phi(R) \leftrightarrow \phi(K) \]

for the periodic wave function, where \( R \) and \( K \) are plane (2D) vectors, respectively, in real space and in reciprocal space.

2.1 The basic formula of the multislice method

The wave function \( \Psi(r) \) of the electron in the crystal obeys the time-independent Schrödinger equation, which is

\[
- \left[ \frac{\hbar^2}{8\pi^2 m} \Delta_r + eU(r) \right] \Psi(r) = \frac{\hbar^2 k^2}{2m} \Psi(r) .
\]  

(1)

The symbols in (1) have their usual meanings (see, e.g., [5]).

For high-energy electrons it is convenient to rewrite the wave function (without any loss of generality) as a modulated plane wave with

\[
\Psi(r) = \varphi(r) e^{2\pi ikz} ,
\]  

(2)

where \( k \) is the length of the electron wave vector. It should be noticed that (2) is different from the usual case where the plane wave is written as a three-dimensional plane wave \( e^{2\pi ik \cdot r} \). However, as we will see, (2) is more convenient for treating beam till and coherent many-beam illumination.

Substituting (2) into (1), and neglecting the term \( \partial^2 \varphi(r)/\partial z^2 \) which mainly represents the back-scattering effect [16, 17], we can finally obtain a first-order differential equation for \( \varphi(r) \) with

\[
\frac{\partial \varphi(r)}{\partial z} = \left[ \frac{i\lambda}{4\pi} \Delta + i\sigma U(r) \right] \varphi(r) ,
\]  

(3)

where \( \sigma \) is the interaction constant, \( \lambda \) the wavelength, and \( \Delta \) the Laplacian operator in the \( x-y \) plane. Most of the current formulations for high-energy electron diffraction can be derived from (3) (for more details, see [7]).
We now suppose that the crystal consists of a series of thin crystal slices with the same thickness and consider the transmission of the electron wave from slice \((n - 1)\) to slice \(n\). If the thickness of the crystal slice is small enough, we can use the projection approximation, under which the potential \(U(r)\) in (3) is assumed to be the average potential with

\[
V_p^\infty(R) = \frac{1}{E} \int_{(n-1)e}^{ne} U(r) \, dz ,
\]

where \(e\) is the thickness of one slice. If \(e\) is smaller than the periodicity in the \(z\)-direction, HOLZ (higher-order Laue zone) effects are included in the multislice method [18].

We therefore obtain the wave function at \(z = ne\) [7] finally as

\[
\varphi_n(R, ne) = e^{\frac{iLe}{\Lambda} + \sigma V_p^\infty(R)} \varphi_{n-1}(R, (n - 1) e) .
\]

Using (5) recursively, the complete wave function at \(z = ne\) is obtained to be

\[
\Psi_n(R, ne) = e^{2\pi ike} \left[ e^{\frac{iLe}{\Lambda} + \sigma V_p^\infty(R)} ... e^{\frac{iLe}{\Lambda} + \sigma V_p^\infty(R)} \varphi_0(R, 0) \right] ,
\]

where \(e^{2\pi ike}\) is a constant phase factor for electrons with a fixed wavelength and therefore can be neglected for all the following cases. Using (2) at \(z = 0\), we can find \(\varphi_0(R, 0) = \Psi_0(R, 0)\). So again the basic multislice formula for calculating the wave function at \(z = ne\) is

\[
\Psi_n(R) = e^{\frac{iLe}{\Lambda} + \sigma V_p^\infty(R)} ... e^{\frac{iLe}{\Lambda} + \sigma V_p^\infty(R)} \Psi_0(R) ,
\]

and therefore using the first-order Zassenhaus expansion [10], we obtain the traditional multislice formula

\[
\Psi_n(R) = e^{\frac{iLe}{\Lambda}} e^{i\sigma V_p^\infty(R)} ... e^{\frac{iLe}{\Lambda}} e^{i\sigma V_p^\infty(R)} \Psi_0(R) ,
\]

where \(\Psi_0(R)\) is the incident wave function (illumination condition).

We now apply (8) for two practical illumination cases:

(i) Tilted plane wave illumination: In this case, \(\Psi_0(R) = \psi_0(K_u) e^{2\pi i K_u \cdot R}\), where \(\psi_0(K_u)\) is the constant complex amplitude of the plane wave and \(K_u\) is the component of the wave vector \(k\) in the \(x-y\) plane (Fig. 1), so we have

\[
\Psi_n^{K_u}(R) = e^{\frac{iLe}{\Lambda}} e^{i\sigma V_p^\infty(R)} ... e^{\frac{iLe}{\Lambda}} e^{i\sigma V_p^\infty(R)} \Psi_0(K_u) e^{2\pi i K_u \cdot R} .
\]

From (9), it can be seen that the effect of beam tilt is included in the incident wave function without any change of the phase grating and the Fresnel propagator.

(ii) Coherent many-beam illumination: If the incident wave on the entrance face of the crystal is not a single plane wave but a coherent superposition of plane waves with

\[
\Psi_0(R) = \sum_{K_u} \psi_0(K_u) e^{2\pi i K_u \cdot R} ,
\]

and assume that the vector \(k\) lies in the plane of the grating, then

\[
\Psi_n^{K_u}(R) = e^{\frac{iLe}{\Lambda}} e^{i\sigma V_p^\infty(R)} ... e^{\frac{iLe}{\Lambda}} e^{i\sigma V_p^\infty(R)} \Psi_0(K_u) e^{2\pi i K_u \cdot R} ,
\]
the total wave function at the exit surface of the crystal is

\[
\Psi_n(R) = e^{i\varkappa_R} \frac{\mathrm{d}z}{\varkappa_R} \frac{\mathrm{d}x}{\varkappa_R} \frac{\mathrm{d}y}{\varkappa_R} \Psi_0(R)
\]

So we can see that for coherent many-beam illumination, the multislice calculation can be performed in two different ways: 1. the incident wave is treated as a whole; 2. separate calculations are made for each incident plane wave and then all the sub-wave functions obtained are added in total.

\[\text{(11)}\]

2.2 The Bloch wave form of the multislice method

Equation (9) is convenient for performing calculations with the real-space algorithm [8, 19, 20], but it is not convenient for performing calculations with the FFT algorithm [11], since the plane wave factor \(e^{i\varkappa_{K_u} \cdot R}\) normally does not have the periodicity of the phase grating.
For the FFT algorithm it will be convenient to give the Bloch wave form of (9). Generally, assuming that the wave function for the \((n - 1)\)-th slice is a Bloch wave in the \(x-y\) plane with
\[
\psi_{n-1}^{K_u}(R) = \phi_{n-1}^{K_u}(R) \cdot e^{2\pi i K_u \cdot R},
\]  
(12)
where \(\phi_{n-1}^{K_u}(R)\) has the same periodicity as the crystal. The wave function for the \(n\)-th slice, according to (8), will be
\[
\psi_{n}^{K_u}(R) = e^{\frac{\iota \vec{d} \cdot \vec{r}}{\iota}} e^{\iota \vec{S} \cdot \vec{Z}(R)} [\phi_{n-1}^{K_u}(R) e^{2\pi i K_u \cdot R}].
\]  
(13)
Performing the Fourier transform over (13) yields
\[
\psi_{n}^{K_u}(K) = \phi_{n}^{K_u}(K) \ast \delta(K - K_u),
\]  
(14)
where \(\ast\) indicates a convolution and
\[
\phi_{n}^{K_u}(K) = P(K + K_u) [Q_n(K) \ast \phi_{n-1}^{K_u}(K)],
\]  
(15)
where \(P(K)\) and \(Q(K)\) are, respectively, the normal Fresnel propagator and the phase grating in reciprocal space. Going back to real space, we have
\[
\psi_{n}^{K_u}(R) = \phi_{n}^{K_u}(R) e^{2\pi i K_u \cdot R},
\]  
(16)
where \(\phi_{n}^{K_u}(R)\) is a periodic function with
\[
\phi_{n}^{K_u}(R) = e^{\frac{\iota \vec{d} \cdot \vec{r}}{\iota}} - i \vec{S} \cdot \vec{V} - \pi \iota e^{\iota \vec{Z} \cdot \vec{V}(R)} \phi_{n-1}^{K_u}(R),
\]  
(17)
where \(\vec{V}\) is the gradient operator in the \(x-y\) plane [8]. So (16) and (17) yield the Bloch wave form of the multislice method and (14) with (15) are the equivalent expressions in reciprocal space. In the Bloch wave form of the multislice method, as we can see, the effects of beam tilt are included in the propagator. Generally, these formulations are suited to the case where the incident wave is a Bloch wave, because the periodic factor of the Bloch wave can combine with the first phase grating without the need of periodic continuation. In single-beam illumination case, the plane wave itself can be regarded as a Bloch wave because any plane wave can be written as a Bloch wave.

The Bloch wave form of the multislice method provides an effective way for calculating coherent many-beam diffraction patterns: 1. first calculating the periodic functions within one unit cell for each beam direction by using (15) and (17), alternatively, with the FFT algorithm; 2. then adding in total all the Bloch waves obtained from (14) (in reciprocal space) or (16) (in real space).

Equations (15) and (17) are different from the traditional multislice formulations [6, 7]. The Fresnel propagator is modified for the case of beam tilt. What is lost in the traditional formulations is the phase factor \(e^{-\iota \vec{S} \cdot \vec{K}_u}\). In case of incoherent illumination and in case of the kind of coherent illumination under which no overlap occurs between diffraction discs, the phase factor \(e^{-\pi \iota \vec{S} \cdot \vec{K}_u}\) acts as a constant. However, for the case of coherent illumination where overlaps between diffraction discs occur [21], the above phase factor is not constant any more. So, generally, in order not to lose the information about the relative phases of the beams corresponding with different incident directions, one has to use (14) to (17) instead of the traditional multislice expression for treating beam tilt [6].

To look for the physical meaning of the phase factor \(e^{-\pi \iota \vec{S} \cdot \vec{K}_u}\), we can write the traditional propagator in more detail, which is
\[
P(K, K_u) = e^{-2\pi i \vec{S} \cdot \vec{K}_u} = e^{-2\pi \iota \iota \frac{1}{2} (K^2 + 2 K_u \cdot K)}.
\]  
(18)
The effects of beam tilt are taken into account by the excitation errors ζ(K). In (18), however, the excitation errors of all the central diffraction beams (K = 0) from different incident directions within the illumination probe are the same, i.e. always zero. If we think that a CBED pattern is a superposition of many sets of normal diffraction patterns — each set of normal diffraction patterns is generated by each incident beam within the probe, then we can see that (18) does not include the relative excitation errors or phase differences between these sets of normal diffraction patterns. In other words, if we look at the central disc and then move away from the origin of the reciprocal lattice of the crystal but keep the Ewald sphere, it can clearly be seen that the phase factor $e^{-i\lambda R K^2}$, which is neglected in the traditional propagator, actually includes the relative excitation errors or phase differences between incident beams from different directions.

3. A Faster Procedure for Calculating Coherent STEM Images

3.1 Coherent STEM image simulations

In the coherent STEM case, the probe function can be expressed as follows:

$$\Psi_0(R, R_p) = \sum_{K_u} \psi_0(K_u) e^{2\pi i K_u \cdot (R - R_p)},$$  \hspace{2cm} (19)

where $R_p$ indicates the probe position and $\psi_0(K_u)$ is a normalised function depending on the illumination conditions of the electron microscope [5].

For each probe position a set of coherent CBED patterns $\psi_n(K, R_p)$ can be obtained by the multislice calculation. A coherent STEM image is the intensity sum of a series of such kind of coherent CBED patterns combined with the detector geometry $G(K)$ [1, 22],

$$I(R_p) = \sum_K |\psi_n(K, R_p)|^2 G(K).$$  \hspace{2cm} (20)

So for coherent STEM image simulations, we need to calculate separately many sets of coherent CBED patterns for each probe position within the area of probe scanning. In other words, the computation time of a coherent STEM image basically depends on how much time is needed to calculate a set of CBED patterns.

Substituting (19) into (8), the basic formula for calculating coherent CBED patterns is

$$\Psi_n(R, R_p) = e^{i\xi_{nd}} e^{i\alpha V_0^g(R)} \cdots e^{\frac{i\xi_{nd}}{4\pi}} e^{i\lambda V_1^g(R + R_p)} \sum_{K_u} \psi_0(K_u) e^{2\pi i K_u \cdot (R - R_p)} \cdot \sum_{K_u} \psi_0(K_u) e^{2\pi i K_u \cdot R_p}.$$  \hspace{2cm} (21)

Equation (21) means that when performing the calculation, the effects of the probe scanning can be included by shifting the object (the phase grating shifting) in the opposite direction of the probe scanning.

3.2 A faster procedure

In real space, supposing $N_xN_y(= n)$ of unit cells are needed to cover a sufficiently large area and in each unit cell the number of sampling points is $N_xN_y(= N)$, the total number of sampling points will be $N_xN_yN_xN_y$ or $nN$. Correspondingly, in reciprocal space this means that $N_xN_y$ reciprocal lattice points are included — each of them having a Brillouin-zone area containing $N_xN_y(= n)$ sampling points (diffraction beams).
For the conventional procedure, the probe function is treated as a whole by using (21). The computer time for each probe position is

\[ T_0 = C[N_u N_b N_x N_y \log_2 (N_u N_b N_x N_y)] \]
\[ = C(nN \log_2 N + nN \log_2 n), \]  
(22)

where \( C \) is a constant [1].

If calculations are separately made for each beam within the illumination aperture by using the Bloch wave form of the multislice method, that is (14) to (17), calculations can be performed only within one unit cell and therefore the total computer time for \( N_u N_v (= n') \) incident beams will be

\[ T' = CN_u N_v (N_x N_y \log_2 N_x N_y) \]
\[ = Cn'N \log_2 N, \]  
(23)

where the number of beams depends on the size of illumination aperture (for a fixed sampling interval). In case where the size of the illumination aperture is equal to or smaller than the first Brillouin zone, that is \( n' \leq n \), \( T' \) is obviously less than \( T_0 \). In other cases, however, the illumination aperture can be larger or much larger, i.e. the case of coherent overlaps between diffraction discs and the shadow imaging case [2], and it is difficult to compare \( T' \) with \( T_0 \) because \( T' \) will be larger than \( T_0 \) when the beam number \( n' \) reaches a certain value.

Fortunately, in such cases we can classify the beams within the coherent probe and combine them into a few artificial Bloch waves,

\[ \Psi_0(R) = \sum_{K_u} \psi_0(K_u) e^{2\pi i K_u \cdot R} \]
\[ = \sum_{k_i} \left[ \sum_{K_u} \psi_0(K_u + k_i) e^{2\pi i K_u \cdot R} \right] e^{2\pi i k_i \cdot R} \]
\[ = \sum_{k_i} \Phi_0^k(R) e^{2\pi i k_i \cdot R}, \]  
(24)

where \( K_u = K_n + k_i \), \( K_n \) is a reciprocal lattice vector and \( k_i \) is the Bloch wave vector restricted in the first Brillouin zone (Fig. 2). Equation (24) means that the incident wave can be thought to be a superposition of a series of Bloch waves and the functions \( \Phi_0^k(R) \) have the same periodicity as the crystal in the \( x - y \) plane. In other words, we can combine the beams outside and inside the first Brillouin zone into a few incident Bloch waves.

![Diagram](image)

Fig. 2. Illustration of an incident beam \( K_u \) outside of the first Brillouin zone being related by a reciprocal lattice vector \( K_n \) to the incident beam \( k_i \) located in the first Brillouin zone.
According to Section 2.2, the Bloch wave form of the multislice method is not only suited to a single-beam illumination but also generally suited to a Bloch wave illumination which may contain a few plane waves related by reciprocal lattice vectors (Fig. 2). So when performing calculations, we can treat each of such incident Bloch waves as a single incident plane wave by using (14) to (17), and therefore get benefits in computing time.

In summary, a new procedure for coherent STEM image simulations can be described as follows: first calculations are separately made for each of the incident Bloch waves $k_i$ with the FFT algorithm and then the obtained sub-wave functions are summed up:

1. In real space

$$\psi_n(R, R_p) = \sum_{k_i} \psi_n^{k_i}(R, R_p) = \sum_{k_i} \phi_n^{k_i}(R, R_p) e^{2\pi ik_i \cdot R},$$

where $k_i$ is located in the first Brillouin zone and

$$\phi_n^{k_i}(R, R_p) = e^{\frac{ik_i A}{4\pi}} \frac{i e^{ik_i \cdot \nabla} \nabla \cdot e^{-ik_i \cdot \nabla_e \nabla_i}}{\epsilon_k \epsilon_{k_i}} e^{\frac{e^{ik_i \cdot R}}{4\pi}} \phi_{0}^{k_i}(R, R_p).$$

(26)

where $\phi_{0}^{k_i}(R)$ is defined by (24).

2. In reciprocal space

$$\psi_n(K, R_p) = \sum_{k_i} \psi_n^{k_i}(K, R_p)$$

$$= \sum_{k_i} \phi_n^{k_i}(K, R_p) \ast \delta(K - k_i),$$

(27)

where

$$\phi_n^{k_i}(K, R_p) = P(K + k_i)$$

$$\times \left[ Q_n(K, R_p) \ast \ldots \left[ P(K + k_i) Q_1(K, R_p) \ast \phi_{0}^{k_i}(K) \right] \ldots \right],$$

(28)

where $\phi_{0}^{k_i}(K)$ is, according to (24), defined by

$$\phi_{0}^{k_i}(K) = \sum_{K_n} \psi_0(K_n + k_i) \delta(K - K_n).$$

(29)

The computing time of the new procedure actually depends on the number $(M_b)$ of Bloch waves which can be created within the incident beams by (24),

$$T = CM_bN \log_2 N.$$

(30)

The computing times of the conventional procedure (22) and the new procedure (30) are therefore compared as follows:

1. When the illumination aperture is smaller than the first Brillouin zone, $M_b$ is the number of beams within the aperture (in this case each beam itself is a Bloch wave), $T$ is obviously less than $T_0$.

2. When the illumination aperture is larger than the first Brillouin zone, $M_b = n$, which is the number of sampling points inside the first Brillouin zone. In this case, (i) if $n > 1$, $T < T_0$, the more unit cells are needed for the periodic continuation in the conventional procedure, the more benefits will be gained by using the new procedure; (ii) if $n = 1$, $T = T_0$. For large unit cell structures or nonperiodic structures, only one unit cell or artificial unit cell is needed to include enough sampling points for the conventional procedure. In this case, the new procedure actually returns to the conventional one: the first Brillouin zone
is so small that it contains only one sampling point and all the illumination beams are combined into a single Bloch wave. (iii) Furthermore, the new procedure will be much faster in such cases where only the intensity distribution along a cross-line in the CBED patterns is needed. For example, the computing time for calculating the intensity distribution along [100] in a ronchigram or shadow image where lattice fringes can appear without moving the probe [2, 5], will be

\[ T^* = CN_s N \log_2 N. \]

Some results of practical computing time for the traditional procedure and the new procedure are shown in Table 1.

3. The new procedure shows its largest advantage when only partly overlap occurs between two diffraction discs. In this case, because the intensities of CBED patterns change with the probe position only in the overlapping regions [5], calculations are needed to be made only for the incident Bloch waves which generate these overlapping regions (Fig. 3).

4. Conclusion

A few convenient expressions of the multislice method are derived from the basic Schrödinger equation. It is found that the effects of beam tilt can be included in the multislice method in two different ways:

| Table 1 |
| Results of practical computing test for the traditional procedure (TP) and the new procedure (NP), i.e. time measured in units of s/20 slices |
| sampling points | \(8 \times 8 \times 32 \times 32\) complete CBED patterns | \(8 \times 8 \times 32 \times 32\) CBED intensity along [100] |
| TP | 25.6 | 25.6 |
| NP | 17.3 | 2.2 |
1. The beam tilt is automatically included in the incident plane wave function without any change of the Fresnel propagator. The new expression provides an easy way to understand the relationship between the total wave function and the sub-wave functions contributed by separate plane-wave illuminations.

2. The beam tilt is included in the propagator. In this case, the multislice method appears in a Bloch wave form. When performing calculations by using these formulations for each coherent incident direction in the STEM case, calculations can be carried out only in one unit cell and the probe scanning can be included by shifting the phase grating, however, all the relative phases should be included. Furthermore, it is shown that if the incident wave is a Bloch wave, it can be treated as a single-beam illumination. This finally leads to a faster procedure for simulating coherent STEM images.

Formulations for carrying out the new procedure are described and advantages of the new procedure are discussed with respect to a few illumination conditions. Generally, the new procedure can be faster than the conventional procedure if more than one unit cell are needed in the periodic extension. For non-periodic structures or structures with very large unit cells (in $x-y$ plane), only one unit cell might be enough; in this case, the new procedure actually returns to the conventional procedure.

Acknowledgements

This text presents research results partly sponsored by the Belgian programme on InterUniversity Poles of Attraction initiated by the Belgian State, Prime Minister’s Office of Science Policy Programming. The scientific responsibility is assumed by the authors. The authors are grateful to Dr. X. B. Zhang and Mr. F. Schallenberg for their help in preparing the manuscript.

References