Ultramicroscopy 14 (1984) 351–356
North-Holland, Amsterdam

BLOCH WAVE HREM

L.D. MARKS

Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

Received 21 May 1984; revised manuscript received 16 July 1984

An analysis of the effects of beam convergence is carried out for high resolution electron microscope imaging, including the changes in diffraction across the illumination aperture by a Bloch wave approach. It is shown that the effect of a small tilt $\boldsymbol{u}$ is to introduce a perturbation $\boldsymbol{u}\cdot\nabla_\rho$ into the projected 2D Schrödinger equation for the Bloch waves. By using perturbation theory (essentially a $\boldsymbol{k}\cdot\boldsymbol{p}$ analysis), it is shown that the dominant effect is a modification of the non-linear convergence envelope term between different Bloch waves. The effects should be larger for defects than for single crystals, and important for thicknesses of more than ~ 50 nm, thicknesses similar to those at which non-linear effects are important. The general trends with structure, elemental composition and incident electron voltage are briefly described.

1. Introduction

Over the past few years considerable success has been achieved in high resolution electron microscopy using the technique of structure imaging to examine local inhomogeneities (e.g. ref. [1]). The majority of this work has been concerned with thin specimens, combining the experimental results with fairly detailed numerical calculations performed by the multislice technique (e.g., refs. [2,3]). Thicker specimens have often been ignored, primarily because the numerical calculations often give only poor agreement with the experimental results (which in some cases may be due to non-linear terms). Progress towards an understanding of thicker specimen has been slow, one of the main reasons being the difficulty in locating errors or invalid approximations inside a "black-box" numerical routine.

Although analytical techniques are well established using Bloch wave techniques, e.g. ref. [4], relatively little work has been performed using these for thicker specimen HREM. Some attempts have recently been made [5–7] to correlate HREM with Bloch waves, but only in the absence of the imaging envelope terms, which is a poor approximation to an operating electron microscope.

In this paper we analyze the effect of beam convergence by including the variations in the diffraction conditions by a Bloch wave expansion, and using perturbation theory to describe these. The required perturbation has the very simple form which is proportional to the momentum operator, which implies that the effects can be large around crystalline defects. (The analysis is essentially a $\boldsymbol{k}\cdot\boldsymbol{p}$ perturbation analysis – see, for instance, ref. [8].) With a few small approximations and simplifications, we find that errors arise in conventional non-linear imaging equations for typical thicknesses of more than ~ 50 nm.

2. Theory

In the framework of a Bloch wave solution for swift electron diffraction, we consider a series solution in the form (e.g. refs. [4,9–12]),

$$\psi(\boldsymbol{r}) = \sum_j C_0^j(\boldsymbol{v}) \exp\left\{-2\pi i\left[z/\lambda - s_j(\boldsymbol{v})z\right]\right\} |j\rangle, \tag{1}$$

where $\boldsymbol{k} = z/\lambda + \boldsymbol{v}$ specifies the incident beam direction, and $|j\rangle$ and $s_j(\boldsymbol{v})$ are respectively the eigenfunctions and eigenvalues of

$$H_0|j\rangle = (8\pi^2/\lambda)\, s_j(\boldsymbol{v})\, |j\rangle, \tag{2}$$

with

$$H_0 = \nabla_\rho^2 + \frac{8\pi^2 me}{h^2} V_\rho(\boldsymbol{r}), \tag{3}$$

$|j\rangle$ a Bloch wave of form

$$|j\rangle = \sum_g C_g(\boldsymbol{v}) \exp[-2\pi i(\boldsymbol{v}+\boldsymbol{g})\cdot\boldsymbol{\rho}], \tag{4}$$

and $V_\rho(\boldsymbol{r})$ the projected potential, with $\boldsymbol{v}$ and $\boldsymbol{\rho}$ respectively reciprocal and real space vectors in the zone axis plane, z down the zone axis, and λ the electron wavelength. Eq. (2) is equivalent to solving a two-dimensional band structure problem [10,12].

Considering now a small change in the incident direction specified by a vector $\boldsymbol{u}$ in the diffraction plane, the change in H_0 takes the form:

$$H_0 \rightarrow H_0 + \boldsymbol{u}\cdot\boldsymbol{H}', \tag{5}$$

where

$$\boldsymbol{H}' = \nabla_{\boldsymbol{\rho}}. \tag{6}$$

Dealing with $\boldsymbol{H}'$ by perturbation theory, we make the standard expansion

$$|j\rangle \rightarrow |j\rangle + \boldsymbol{u}\cdot{\sum_m}' \boldsymbol{a}_m^j |m\rangle + \ldots, \tag{7}$$

$$s_j(\boldsymbol{v}) \rightarrow s_j(\boldsymbol{v}) + \boldsymbol{u}\cdot\boldsymbol{s}_j^1(\boldsymbol{v}) + \sum_{lm} u_l u_m s_j^2(\boldsymbol{v})_{lm} + \ldots, \tag{8}$$

and assuming (for simplicity) that non-degenerate perturbation theory can be used, we have the standard results

$$\boldsymbol{a}_m^j = (\lambda/8\pi^2)\langle m|\boldsymbol{H}'|j\rangle / (s_j(\boldsymbol{v}) - s_m(\boldsymbol{v})), \tag{9}$$

$$(8\pi^2/\lambda)\,\boldsymbol{s}_j^1(\boldsymbol{v}) = \langle j|\boldsymbol{H}'|j\rangle, \tag{10}$$

$$(8\pi^2/\lambda)\, s_j^2(\boldsymbol{v})_{lm} = {\sum_n}' \langle n|H_l'|j\rangle\langle j|H_m'|n\rangle / [s_j(\boldsymbol{v}) - s_n(\boldsymbol{v})]. \tag{11}$$

Before proceeding further, we should point out the physical significance of some of the terms in eqs. (5)–(11). The perturbation $\boldsymbol{H}'$ is proportional to the momentum operator, so $\boldsymbol{s}_j^1(\boldsymbol{v})$ is proportional to the expectation of the transverse momentum of the jth Bloch wave in the 2D potential of eq. (2). Assuming correct orientation of the crystal down a zone axis, $\boldsymbol{s}_j^2(\boldsymbol{v}) = \boldsymbol{0}$. We also note that both $\boldsymbol{a}_m^j$ and the terms in the sum in eq. (11) are inversely proportional to the difference $s_j(\boldsymbol{v}) - s_m(\boldsymbol{v})$. Hence the main interactions arise from adjacent branches of the dispersion surface. Finally, since $\boldsymbol{H}'$ is a gradient operator, it follows that the effects of illumination angle are most important whenever the wavefunction is rapidly varying. This implies larger effects at any crystal defects than in the bulk crystal. The effect of the operator also increases with larger $\boldsymbol{g}$ values, and hence becomes more important with higher resolution (assuming the convergence is constant). This is important for current high voltage machines which are primarily limited by the chromatic aberration terms.

The expansion of $|j\rangle$ and $s_j(\boldsymbol{v})$ to first and second order respectively will be valid for most experimental cases where the convergence is small compared to the first Brillouin zone size, the latter being the natural unit of scale in reciprocal space in a Bloch wave approach. We only require relatively small changes in $|j\rangle$ and $s_j(\boldsymbol{v})$, which can lead to large changes in $\psi(\boldsymbol{r})$ if z is large. Writing out the expansion of $\psi(\boldsymbol{r}, \boldsymbol{u})$ for variations in the incident directions, we have

$$\psi(\boldsymbol{r}, \boldsymbol{u}) = \sum_j \left[C_0^j(\boldsymbol{v}) + {\sum_m}' C_0^m(\boldsymbol{v})\, \boldsymbol{u}\cdot\boldsymbol{a}_m^j \right] \times \exp\left\{ -2\pi i \left[z/\lambda + \boldsymbol{u}\cdot\boldsymbol{r} - z\left(s_j(\boldsymbol{v}) + \boldsymbol{u}\cdot\boldsymbol{s}_j^1(\boldsymbol{v}) + \sum_{lm} u_l u_m\, s_j^2(\boldsymbol{v})_{lm} \right) \right] \right\} \left(|j\rangle + {\sum_m}' \boldsymbol{u}\cdot\boldsymbol{a}_m^j |m\rangle \right). \tag{12}$$

It is appropriate to briefly discuss the first-order correction to the amplitude, $\boldsymbol{u}\cdot\sum_m' C_0^m(\boldsymbol{v})\boldsymbol{a}_m^j$, and to the Bloch wave, $\boldsymbol{u}\cdot\sum_m' \boldsymbol{a}_m^j|m\rangle$. If either or both of these are important in the final result, the conventional reciprocal space integration will be invalid *even* for a very thin crystal. For the moment we will assume that the crystal has a center of symmetry and is accurately aligned on a zone

axis. For a symmetric Bloch wave, both of these first-order terms have the wrong parity and must be zero. The only case where they can be non-zero is for an asymmetric Bloch wave. However, both first-order terms have odd parities, and must therefore be paired with another odd term in order to yield any effect in the final (symmetric) integration over the convergence. The only other odd terms are other first-order corrections, or terms due to the spherical abberration and defocus phase shift, χ, of form $\sin(\boldsymbol{u}\cdot\nabla\chi)$. Both of these are inherently small for small $\boldsymbol{u}$, and the final contribution is of order u^2, i.e. essentially a second-order correction.

Therefore, as a rule the effect of these two first-order terms is small, and it is justifiable to ignore them, although there is an important exception which is not covered by the above analysis. In many crystals certain diffracted beams are both kinematically and dynamically forbidden on the zone axis, but become dynamically allowed off the axis. (They would show Gjonnes–Moodie lines in a convergent beam pattern.) The integration over diffraction conditions will lead to these spacings appearing in the image, whilst they would be completely absent in a simple reciprocal space integration. We note that similar effects can also occur at an extinction contour.

The last two paragraphs have assumed that the crystal has a center of symmetry and is correctly aligned. We note that the above arguments, that the first-order amplitude and wave terms are small, may well break down if these conditions are not met.

Continuing the analysis, the important perturbations (with the qualifications detailed above) are $s_j^1(\boldsymbol{v})$ and $s_j^2(\boldsymbol{v})_{lm}$, since these lead to large changes if z is large. Neglecting the other terms we can approximate:

$$\psi(\boldsymbol{r},\boldsymbol{u}) \approx \sum_j C_0^j(\boldsymbol{v}) \exp\Big\{-2\pi\mathrm{i}\Big[z/\lambda + \boldsymbol{u}\cdot\boldsymbol{r} - z\Big(s_j(\boldsymbol{v}) + \boldsymbol{u}\cdot\boldsymbol{s}_j^2(\boldsymbol{v}) + \sum_{lm} u_l u_m s_j^2(\boldsymbol{v})_{lm}\Big)\Big]\Big\}\,|j\rangle. \tag{13}$$

The exact variation of $s_j(\boldsymbol{v})$ will in general depend upon the specimen employed and the accelerating voltage of the microscope. Therefore, rather than becoming involved in one particular experimental set-up, we choose here to look in a more general fashion at the effects and extract the trends. Assuming a correct orientation down a zone axis, we can omit $s_j^1(\boldsymbol{v})$ and estimate $s_j^2(\boldsymbol{v})_{lm}$. (We note that tilting the crystal off the zone axis has a number of large disadvantages in terms of the image localization, as discussed elsewhere [12].) Taking the dispersion surface curvature to be of the same order as for a free electron sphere, a reasonable approximation is

$$s_j^2(\boldsymbol{v})_{lm} = \lambda\alpha_j\delta_{lm}, \tag{14}$$

where $|\alpha_j| \sim 1/2$, and α_j can be positive or negative.

We now carry out the analytical imaging process using the approximate form of $\psi(\boldsymbol{r},\boldsymbol{u})$. For convenience we will separate the series terms in $\psi(\boldsymbol{r},\boldsymbol{u})$,

$$\psi(\boldsymbol{r},\boldsymbol{u}) = \sum_j \psi_j(\boldsymbol{r}) \exp\big[-2\pi\mathrm{i}\big(\boldsymbol{u}\cdot\boldsymbol{r} - \lambda\alpha_j u^2 z\big)\big], \tag{15}$$

and separate the image intensity due to different cross-terms between the Bloch waves:

$$I(\boldsymbol{r}) = \sum_{lm} I_{lm}(\boldsymbol{r}). \tag{16}$$

Writing $\mathscr{F}$ for a Fourier transform operator, with

$$\Psi_j(\boldsymbol{v}) = \mathscr{F}\psi_j(\boldsymbol{r}), \tag{17}$$

and performing the standard incoherent integral for convergence and chromatic terms (e.g. refs. [13–16]), we find

$$I_{lm}(\boldsymbol{r}) = \mathscr{F}^{-1}\int \Psi_l^*(\boldsymbol{w})\,\Psi_m(\boldsymbol{v}-\boldsymbol{w}) \times T_{lm}^{\mathrm{conv}}(\boldsymbol{w},\boldsymbol{v}-\boldsymbol{w},z) \times \exp\{-\mathrm{i}[\chi(\boldsymbol{v}-\boldsymbol{w})-\chi(\boldsymbol{w})]\} \times E(\boldsymbol{w},\boldsymbol{v}-\boldsymbol{w})\,\mathrm{d}^2\boldsymbol{w}, \tag{18}$$

where $E(\boldsymbol{w},\boldsymbol{v}-\boldsymbol{w})$ is the standard non-linear envelope term which includes *both* convergence and

energy spread terms, and $T_{lm}^{\mathrm{conv}}(\boldsymbol{w}, \boldsymbol{v}-\boldsymbol{w}, z)$ is a thickness-dependent envelope correction term. This latter term includes the effects of changes in the diffraction across the illumination aperture, and takes the form

$$T_{lm}^{\mathrm{conv}}(\boldsymbol{w}, \boldsymbol{v}-\boldsymbol{w}, z) = (1+A^2)^{-1/2} \exp\left\{-\pi^2 \boldsymbol{q}^2 \left[\frac{(1+\mathrm{i}A)}{1+A^2} - 1\right]\right\}, \tag{19}$$

where

$$A = 2\pi\lambda z(\alpha_l - \alpha_m)a^2, \tag{20}$$

$$\boldsymbol{q} = a[\nabla\chi(\boldsymbol{w}) - \nabla\chi(\boldsymbol{v}-\boldsymbol{w})], \tag{21}$$

and the convergence spread has the form $(\pi a^2)^{-1/2} \exp(-u^2/a^2)$; i.e., the convergence contribution to $E(\boldsymbol{w}, \boldsymbol{v}-\boldsymbol{w})$ is $\exp(-\pi^2 q^2)$. Plots of the real part of the correction term as a function of A for several different values of $B = \exp(-\pi^2 q^2)$ are shown in fig. 1 (the imaginary component is negligibly small). Also shown on the axis of the figure are the corresponding thicknesses in nm, assuming a convergence half-height of 1 mRad at 200 kV and $(\alpha_l - \alpha_m) = 1$.

The form of the correction function is a little surprising. Spacings for which there is minimal convergence damping become more strongly damped as the thickness increases. Heavily damped spacings behave in the opposite sense, and are reinforced at larger thicknesses. In many respects the correction term extends the resolution of the microscope a little. Clearly, the effects are important for thicker specimens, from the analysis, beyond about 50 nm.

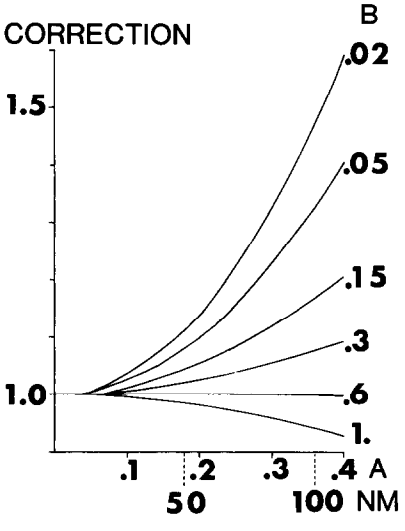

Fig. 1. Plots of the real part of the correction term for different values of B, the damping, as described in the text. Note that the largest nett effect is for undamped ($B = 1.0$) terms.

3. Discussion

The analysis herein has permitted some insight into how convergence (and crystal tilt) can affect a HREM image for a thicker specimen. The effect changes the damping of fringe terms between dispersion surface branches with a different curvature, and can be expected to affect the images for thicknesses larger than ~ 50 nm. The modification will be larger around any form of crystalline defect, coming from the gradient operator. We note that this is an effect *beyond* that of a non-linear envelope term, and it is a moot point to note that some of the "commercial" programs do not even include non-linear envelope terms. However, it is not the only thick specimen term, and phonon and plasmon scattering may need to be considered as well.

It is useful to consider the general trends with materials and incident electron energy. Using the solid-state interpretation of eqs. (6)–(7) [10,12], the variation of $s_j(\boldsymbol{v})$ can be considered in terms of the band structure of a two-dimensional potential. The more tightly bound the potential is, the smaller is the variation and hence the convergence effects. For a given material, the binding strength increases with the size of the projected unit cell. With a given structure, the binding depends upon both the strength and tightness of the atomic scattering. For instance, gold yields a tighter binding than aluminum. In general, the stronger the diffraction, the tighter is the potential. Finally, the binding strength increases with the electron energy, favoring the use of high voltage machines (a relativistic effect).

As a final point, it should be remembered that the role of diffraction changes upon the conver-

gence can be *experimentally* estimated from a selected area diffraction pattern of the area of interest (not from a microdiffraction or convergent beam pattern), with no changes in the illumination conditions, of course.

4. Conclusions

Variations in the dynamical diffraction due to the tilt across the convergence aperture can be expected to require proper treatment beyond that of a non-linear transfer function for typical thicknesses of more than 50 nm, or around any crystal defects where the wavefunction is rapidly varying.

Acknowledgments

The author would like to thank Dr. G.J. Wood for useful discussion. This work was supported by a Department of Energy Grant No. DE-AC02-76ER02995. The manuscript was refereed by Dr. K. Kambe.

References

[1] D.J. Smith, Helv. Phys. Acta 56 (1983) 464.
[2] A. de Crecy and A. Bourret, Phil. Mag. A47 (1983) 245.
[3] L.D. Marks, Surface Sci. 139 (1984) 281.
[4] P.B. Hirsch, A. Howie, R.B. Nicholson, D.W. Pashley and M.J. Whelan, Electron Microscopy of Thin Crystals (Butterworths, London, 1965).
[5] F. Fujimoto, Phys. Status Solidi (a) 45 (1978) 99.
[6] P. Pirouz, Acta Cryst. A37 (1981) 465.
[7] K. Kambe, Ultramicroscopy 10 (1982) 223.
[8] C. Kittel, Quantum Theory of Solids (Wiley, New York, 1963).
[9] K. Fujwara, J. Phys. Soc. Japan 17 (1961) 2226.
[10] M.V. Berry, J. Phys. C4 (1971) 697.
[11] D. Gratias and R. Porter, Acta Cryst A39 (1983) 576, and the references therein.
[12] L.D. Marks, Acta Cryst., submitted.
[13] M.A. O'Keefe, in: Proc. 37th Annual EMSA Meeting, San Antonio, TX, 1979 (Claitor's, Baton Rouge, LA, 1979) p. 556.
[14] W.O. Saxton, J. Microsc. Spectrosc. Electron. 5 (1980) 661.
[15] K. Ishizuka, Ultramicroscopy 5 (1980) 661.
[16] L.D. Marks, Ultramicroscopy 12 (1983–84) 237.