

LATTICE IMAGING IN TRANSMISSION ELECTRON MICROSCOPY

Miroslav Karlík

Department of Materials, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Trojanova 13, 120 00 Prague 2, Czech Republic,
E-mail: karlik@kmat.fjfi.cvut.cz

Abstract

With the resolution becoming sufficient to reveal individual atoms, high-resolution electron microscopy (HREM) can now compete with X-ray and neutron methods to determine quantitatively atomic structures of materials, with the advantage of being applicable to non-periodic objects such as crystal defects. An introduction to the theory and practical aspects of HREM is given. Principles of other lattice imaging techniques in transmission electron microscopy – electron holography and Z-contrast imaging are also described.

Keywords:

High-resolution electron microscopy, electron-specimen interactions, electron diffraction, phase contrast, contrast transfer, electron holography, Z-contrast imaging

1. Introduction

The last decades are characterized by an evolution from macro- to micro- and more recently to nanotechnology. Examples are numerous, such as nanoparticles, nanotubes, quantum transistors, layered superconducting and magnetic materials, etc. Since many material properties are strongly connected to the electronic structure, which in turn is considerably dependent on the atomic positions, it is often essential for the materials science to determine atom positions down to a very high precision. Classical X-ray and neutron techniques fail for this task, because of a non-periodic character of nanostructures. Only fast electrons are scattered sufficiently strongly with matter to provide local information at the atomic scale.

One of the most commonly used high-resolution techniques in transmission electron microscopy (TEM) is that of (bright-field) phase-contrast imaging. This relies on the interference between beams scattered by a specimen and is usually performed under parallel beam illumination conditions. The main and significant disadvantage of this lattice imaging method is in the difficulty of image interpretation in terms of the atomic structure of the specimen. The interpretation is carried out using computer simulations requiring input of a specimen structure model, specimen thickness and microscope parameters.

Unfortunately, even the best electron microscopes are hampered by the fact that only the intensity (i.e. the square of the amplitude), of the electron wave can be recorded on the photographic film or CCD-camera and an essential part of the electron wave, the phase information, is lost. Holographic recording overcome this problem. The (aberrated) image wave is superimposed with an unscattered plane reference wave resulting in an interference pattern - electron hologram. After acquisition and transfer to a computer system, amplitude and phase of the electron wave are recon-

structed using the laws of Fourier optics by sophisticated image processing. Electron – specimen interaction is then simulated in the same manner as in the case of conventional phase-contrast imaging.

On the other hand the Z-contrast technique provides directly interpretable images - maps of scattering power of the specimen. Allowing incoherent imaging of materials, it represents a new approach to high-resolution electron microscopy. The Z-contrast image is obtained by scanning an electron probe of atomic dimensions across the specimen and collecting electrons scattered to high angles. Simultaneously, spectroscopic techniques can also be used to supplement the image, giving information on atomic-resolution chemical analysis and/or local electronic band structure. The resolution of the technique is determined by the size of the electron probe.

This paper gives an introduction to the theory and practical aspects of high-resolution electron microscopy. Principles of electron holography and Z-contrast imaging are also described.

2. Electron – crystal interactions

In a transmission electron microscope, a sample in the form of a thin foil is irradiated by electrons having energy of the order of hundreds of keV. In the interior of the crystal the electrons are either undeviated, scattered, or reflected (Fig. 1).

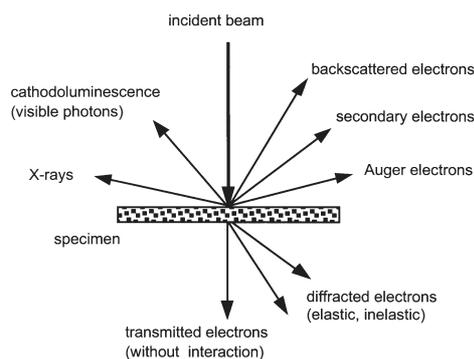


Fig. 1. Interaction of an electron beam with a thin foil

Electron scattering may be either elastic or inelastic. In the case of elastic scattering the electrons interact with the electrostatic potential of atomic nuclei. This potential deviates the trajectory of incident electrons without any appreciable energy loss. In fact a small loss occurs since there is a change in momentum. However, because of the disparity in mass of the scattered electron and the atom the loss is too small ($\Delta E/E \sim 10^{-9}$ at aperture angles used in TEM) to affect the coherency of the beam.



In the inelastic case energy of the incident electron may be transferred to internal degrees of freedom in the atom or specimen in several ways. This transfer may cause excitation or ionization of the bound electrons, excitations of free electrons, lattice vibrations and possibly heating or radiation damage of the specimen. The most common interactions are those with the electrons in the crystal. In this case the energy loss ΔE is important, because the interacting particles have the same mass m . The fraction of energy ΔE is small as compared to the incident energy E . It is this primary process of excitation which is used for the Electron Energy Loss Spectrometry (EELS). This kind of spectrometry as well as different secondary processes of subsequent desexcitations of the target – X-ray emission, Auger electrons emission, cathodoluminescence etc. – permit to link the structural aspect of the specimen with the information about its chemical nature, if the microscope is equipped with appropriate detectors.

2. 1. Schrödinger equation and the crystal potential

It seems obvious that the problem of the diffraction of fast electrons should require the Dirac equation. However, the role of the spin in the interaction is negligible (about one percent [1,2]) and it is therefore sufficient to use the Schrödinger equation:

$$\left[\frac{-\hbar^2}{2m} \Delta + V(\vec{r}) \right] \psi(\vec{r}) = E\psi(\vec{r}) \quad (1)$$

where E is the total energy of the electron and $V(\vec{r})$ its potential energy inside the crystal. For fast electrons ($E \geq 50$ keV), it is necessary to take into account the relativistic effects associated with their motion. The solution of the Schrödinger equation in the crystal, where the moving electron is subjected to the action of a variable potential, is in general very complicated. In case of transmission electron microscopy it is the very high energy of incident electrons ($\sim 10^5$ eV) with respect to the crystal potential (~ 10 eV), which enables to simplify the problem. Due to the very high energy of the incident beam the isotropy of the space is broken: the majority of electrons propagate in the direction of the z axis. The space can be divided into the axial direction (z) and two radial directions (x, y). After this separation of variables it is possible to modify the equation (1) by introducing:

- i) forward (small angle) scattering approximation
- ii) Glauber (projected potential) approximation

The solution of the Schrödinger equation by the multislice method is based on these two approximations. Another approach, often used in the TEM for the solution of the Schrödinger equation is the Bloch wave method, introduced by Bethe in 1928 [3]. Unfortunately, this method can only be used for perfect crystals but not for the crystals containing grain boundaries, precipitates or other faults. The Bloch waves theory is described in detail in [4].

i) Forward scattering approximation

By introducing the relation between the wave vector \vec{k} of the electron beam and corresponding accelerating potential $E = \hbar^2 k^2 / 2m$ to the equation (1) we obtain:

$$\left[\Delta + k^2 \right] \psi(\vec{r}) = \frac{2m}{\hbar^2} V(\vec{r}) \psi(\vec{r}) \quad (2)$$

Owing to the high energy of incident electrons it is possible to consider the action of the crystal potential as a small perturbation resulting only in a modulation $\phi(\vec{r})$ of the primary electron wave. We are seeking a wave function $\psi(\vec{r})$ in the form:

$$\psi(\vec{r}) = \exp(i\vec{k} \cdot \vec{r}) \phi(\vec{r}) \quad (3)$$

Supposing the expression for the modulation $\phi(\vec{r})$, it is possible to write:

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + 2ik_x \frac{\partial}{\partial x} + 2ik_y \frac{\partial}{\partial y} + 2ik_z \frac{\partial}{\partial z} - \frac{2m}{\hbar^2} V(\vec{r}) \right] \cdot \phi(\vec{r}) = 0$$

In this equation the term $\left| \frac{\partial^2 \phi(\vec{r})}{\partial z^2} \right| \ll \left| 2ik_z \frac{\partial \phi(\vec{r})}{\partial z} \right|$ may be neglected since the derivative $\frac{\partial \phi(\vec{r})}{\partial z}$ changes only slowly with z . In other words it is possible to neglect backscattered electrons. The equation (1) thus acquires the form of the forward scattering approximation:

$$\frac{\partial \phi(\vec{r})}{\partial z} = \frac{i}{2k_z} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2ik_x \frac{\partial}{\partial x} + 2ik_y \frac{\partial}{\partial y} - \frac{2m}{\hbar^2} V(\vec{r}) \right] \phi(\vec{r})$$

Two terms in the square brackets may be considered as independent. The first term:

$$A = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2ik_x \frac{\partial}{\partial x} + 2ik_y \frac{\partial}{\partial y}$$

represents the propagation of the electron wave, while the second one represents the action of the crystal potential. Therefore:

$$\frac{\partial \phi(\vec{r})}{\partial z} = \frac{i}{2k_z} \left[A - \frac{2m}{\hbar^2} V(\vec{r}) \right] \phi(\vec{r}) \quad (4)$$

The physical meaning of these two terms becomes obvious from the following: if in the equation (4) we neglect the term $V(\vec{r})$ and if we put $k_x = k_y = 0$ and $k = k_z$ (projection of k to the directions x and y is very small), we obtain:

$$\frac{\partial \phi(\vec{r})}{\partial z} = \frac{i}{2k} \frac{\partial^2 \phi(\vec{r})}{\partial x^2} + \frac{i}{2k} \frac{\partial^2 \phi(\vec{r})}{\partial y^2} \quad (5)$$

This is a 2D equation of heat conduction. Its solution in the point (x, y, z) has a form:

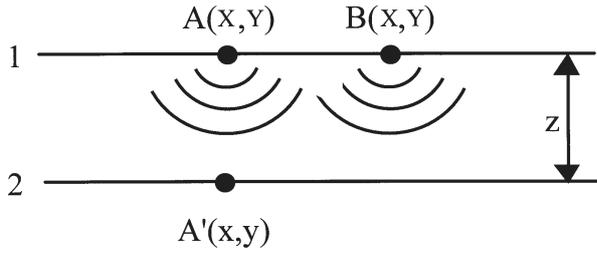


Fig. 2. Propagation of an electron wave

$$\varphi(x, y, z) = \frac{k}{4\pi z} \exp\left(\frac{ik}{2z}(x^2 + y^2)\right)$$

In Fig. 2 A and B there are two point sources of the waves (Huyghens principle) at the level of the plane 1. To find the result of the interference in the point A' at the level of the plane 2 in the distance z , it is necessary to integrate contributions of A, B and all other point sources in the plane 1, represented by the function $\varphi(X, Y, 0)$. Thus we obtain an expression:

$$\varphi(x, y, z) = \frac{k}{4\pi z} \iint \varphi(X, Y, 0) \exp\left[\frac{ik}{2z}((x-X)^2 + (y-Y)^2)\right] dXdY$$

having the form of the convolution integral. Then it is possible to write:

$$\varphi(x, y, z) = \varphi(X, Y, 0) * \left\{ \frac{k}{4\pi z} \exp\left[\frac{ik}{2z}(x^2 + y^2)\right] \right\} \quad (6)$$

where the term in the complex brackets is the Fresnel propagator, describing the propagation of the waves from the plane 1. The result of their interference in the plane 2 is given by the convolution.

ii) Glauber approximation

Neglecting the term A, equation (4) simplifies to the expression:

$$\frac{\partial \varphi(\vec{r})}{\partial z} = -\frac{i}{2k_z} \frac{2m}{\hbar^2} V(\vec{r}) \varphi(\vec{r}) \quad (7)$$

from where we obtain ($k_z \cong k$):

$$\frac{\partial \varphi(\vec{r})}{\partial z} = -\frac{i}{2k} \frac{2m}{\hbar^2} V(\vec{r}) \varphi(\vec{r})$$

and after integration

$$\ln \frac{\partial \varphi(\vec{r})}{\varphi_0(\vec{r})} = -\frac{i}{\hbar v} \int_0^z V(\vec{r}) dz = -\frac{i}{\hbar v} \int_0^z V(xy, z) dz$$

where the constant before the integral was derived from the de Broglie relation $\lambda = h/mv$ (v is the velocity of electrons) and relation $k = 2\pi/\lambda$. For $\varphi(\vec{r})$ we obtain Glauber approximation [5]:

$$\varphi(\vec{r}) = \varphi_0 \exp\left[-\frac{i}{\hbar v} \int_0^z V(xy, z) dz\right] = \varphi_0 \exp(i\chi) \quad (8)$$

and taking $\varphi_0 = 1$ before the interaction of electrons with the crystal, the original expression (3) becomes:

$$\psi(\vec{r}) \cong \exp(i\vec{k} \cdot \vec{r}) [\exp(i\chi)]$$

from which it follows that action of the crystal potential results in a phase shift of the incident electron wave. As the potential energy of electron in the interior of the crystal is negative, the sign of the phase shift is positive.

2. 2. Multislice method

Even in the forward scattering approximation represented by the expression (4), the Schrödinger equation is difficult to be solved. In 1957, Cowley and Moodie [6] proposed an elegant solution, which became much appreciated namely after appearance of computers. Their method consists in separating the equation (4) into two expressions (5) and (7) and in the alternate application of them in thin slices as far as the electron wave pass the whole thickness of the crystal.

Figure 3 represents the crystal in the multislice approximation. The crystal is cut in the direction perpendicular to the incident electron beam into slices of the thickness of the order of one atomic layer. For each of the slices the crystal potential is projected to a plane. The electrons propagate in the "layers" of vacuum on very small

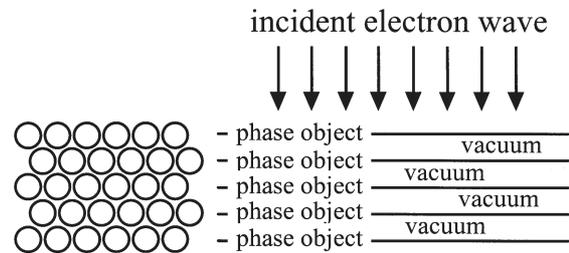


Fig. 3. Representation of the crystal in the multislice approximation

distances and under very small angles. Subsequently they are diffracted on the planes of the potential projected on very small distances and influencing the wavefunction only slightly.

If we adopt Fresnel approximation in which a spherical wavefront is approximated by a parabolic one, we obtain for the wavefunction emerging from the n -th slice:

$$\psi(x, y)_n \approx \frac{\exp(ik'\epsilon)}{\epsilon} [\psi_{in}(X, Y)_{n-1} \exp(i\sigma V_p(X, Y)_n)] * \exp\left[\frac{ik'}{2\epsilon}(X^2 + Y^2)\right]$$

This is the basic formula of the multislice method.



During calculation of the wavefunction two effects alternate. Firstly the phase of the incident electron wave at the entrance to the given slice is shifted due to the action of the projected potential. The phase shift is given by the integral over the thickness ε . Subsequent convolution represents the propagation of the wave in a slice of the vacuum of the same thickness ε . If ε tends to zero, the result of the multislice simulation tends to the exact solution of the Schrödinger equation. As we are dealing with the slices of the finite thickness, it is necessary to determine the largest thickness for which the solution of the Schrödinger equation is still acceptable. Ishizuka and Uyeda [7] give the following condition: $\varepsilon \leq kd^2$, where k is the wave vector of the electrons and d is the distance on which the potential does not change significantly. As follows from this equation, if k is of the order of $2.5 \cdot 10^{12} \text{ m}^{-1}$ and $d \sim 10^{-11} \text{ m}$, ε should be lower or equal to $2,5 \text{ \AA}$! In fact this condition is stronger than shows the practice.

The chart of the simulation corresponding to the EMS code written by Stadelmann [8] is shown in Fig. 4. The incident wave is firstly multiplied by the phase grating, representing the action of the projected potential $V_p(x,y)$. Subsequently a fast Fourier transform (FFT) of this product is performed, in order to calculate easily the product of con-

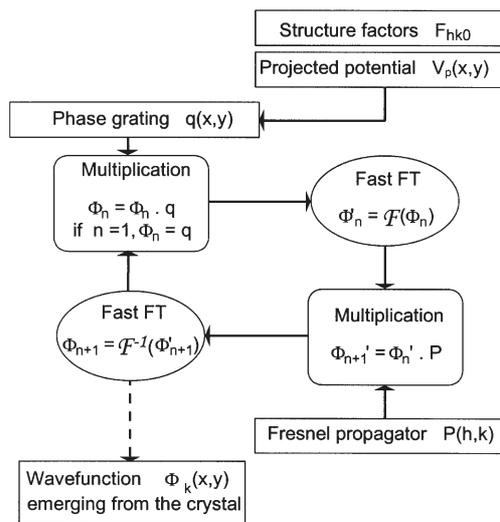


Fig. 4. Multislice simulation calculation chart

volution (multiplication in the Fourier space) with the Fresnel propagator. After the inverse FFT the calculation either continues or ends and we obtain a wavefunction for a given thickness of the crystal.

3. Formation and transfer of the contrast by TEM

A schematic configuration of a TEM column is in Fig. 5. Electrons, emitted by an electron gun pass through accelerator to the illumination system composed of two or more condenser magnetic lenses. After the interaction with the specimen the electrons enter the imaging system of the mi-

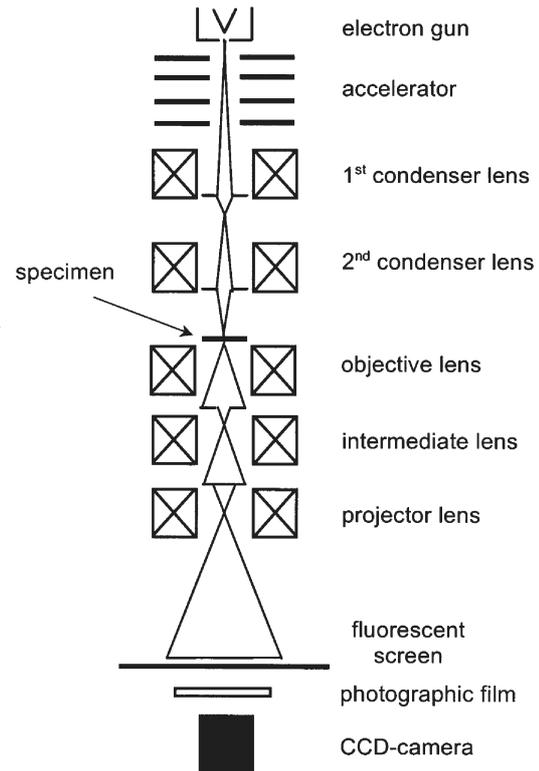


Fig. 5. TEM column

croscope, composed of the objective, intermediate and projector magnetic lenses. The image is visualized on a fluorescent screen and recorded on a photographic film or on a CCD-camera.

Objective lens is the most important lens of TEM, because its aberrations limit the resolution of the microscope. An electron diffraction pattern is formed in its back focal plane. A removable aperture situated in this plane is used to select different electron beams to form different images, thus manipulating the image contrast. In the “classical” diffraction contrast imaging (Fig. 6a) we use a small objective aperture to select only one beam – transmitted (T) or diffracted (D) – to obtain the well-known bright-field or

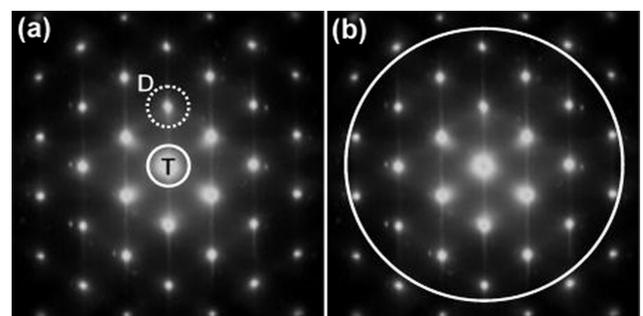


Fig. 6. The size and the position of the objective aperture for diffraction contrast (a) and phase interference contrast (b)

dark-field images of the specimen microstructure. The image is carried only by one beam through the whole imaging

system of the microscope. On the other hand phase-contrast lattice imaging relies on the interference between the transmitted and several diffracted beams and so we use a much larger objective aperture (Fig. 6b).

3.1. Transfer of the contrast by the optical system

In order to understand the relation between the wavefunction at the exit face of the crystal and its image, it is necessary to remind the basic principles of the image formation under coherent illumination [9-11].

Impulse response and transfer function

A system S assigns to the object defined by a function $E(\vec{r})$ an image $I(\vec{r}') = S(E(\vec{r}))$. If the system is linear, it is possible to decompose the object $E(\vec{r})$ to a sum of elementary contributions $e_i(\vec{r})$

$$E(\vec{r}) = \sum c_i e_i(\vec{r})$$

having a weight c_i and to leave the system to act on each of these elements separately. The image $I(\vec{r}')$ of the object is then given by a sum of images of the elementary contributions $S(e_i(\vec{r}))$:

$$I(\vec{r}') = S(E(\vec{r})) = \sum c_i S(e_i(\vec{r}))$$

If the function entering to the system is an elementary impulse (Dirac function δ), we obtain:

$$E(\vec{r}) = \int_{-\infty}^{\infty} E(\vec{r}_1) \delta(\vec{r} - \vec{r}_1) d\vec{r}_1, \text{ and}$$

$$I(\vec{r}') = S(E(\vec{r})) = \int_{-\infty}^{\infty} E(\vec{r}_1) \delta(\vec{r}' - \vec{r}_1) d\vec{r}_1 = \int_{-\infty}^{\infty} E(\vec{r}_1) h(\vec{r}', \vec{r}_1) d\vec{r}_1$$

where $h(\vec{r}', \vec{r}_1)$, the image of the point $\delta(\vec{r} - \vec{r}_1)$, is the impulse response function.

In consequence a linear optical system is fully characterized by point sources in the object plane and by the impulse response function, which describes how the image of any point is influenced by the system. Furthermore, if the system is invariant in the space (isoplanar), its impulse response function $h(\vec{r}', \vec{r}_1)$ depends only on distances $(\vec{r}' - \vec{r}_1)$. In other words, the image of a point source (impulse response to the Dirac function δ) change only its position and does not change either its shape or its intensity, if this point is moved in the object plane. Then it is possible to put $h(\vec{r}', \vec{r}_1) = h(\vec{r}' - \vec{r}_1)$ and simplify the preceding expression, which takes the form of a convolution integral:

$$I(\vec{r}') = \int_{-\infty}^{\infty} E(\vec{r}_1) h(\vec{r}' - \vec{r}_1) = E(\vec{r}) * h(\vec{r}') \quad (9)$$

If we apply a Fourier transform on this expression, we obtain a simple equation which links the spectra $\tilde{I}(\vec{q})$ and $\tilde{E}(\vec{q})$ of the image $I(\vec{r}')$ and object $E(\vec{r})$:

$$\tilde{I}(\vec{q}) = \tilde{E}(\vec{q}) \cdot \tilde{h}(\vec{q})$$

Fourier transform of the impulse response function in this equation,

$$\tilde{h}(\vec{q}) = \int_{-\infty}^{\infty} h(\vec{r}) \exp(i\vec{q}\vec{r}) d\vec{r} = T(\vec{q}) \quad (10)$$

is the transfer function of the system which acts in the domain of spatial frequencies \vec{q} .

Scherzer theory [12] permits to describe the transmission electron microscope as a linear, space invariant system without noise, which is characterized by a transfer function:

$$T(q, \Delta z) = \exp(i\chi(\vec{q}, \Delta z)) = \cos \chi(\vec{q}, \Delta z) + i \sin \chi(\vec{q}, \Delta z) \quad (11)$$

where Δz is the defocus of the objective lens. The real term describes the transfer of the amplitude, while the imaginary term describes the transfer of the phase contrast (this term is often abbreviated as CTF: "Contrast Transfer Function"). Here $\chi(\vec{q}, \Delta z)$ represents the phase shift of the wave as a function of the defocus Δz and the spatial frequency \vec{q} . Spatial frequency in the argument of the transfer function is often replaced by the angle of deviation from the optical axis, which is linked with the spatial frequency by the relation: $\theta = \lambda|\vec{q}|$, where $\theta = 2\theta_{\text{Bragg}}$.

3.2. Image forming by a lens without aberrations

Image formation in a high-resolution electron microscope is an interference phenomenon. A parallel, coherent incident beam is diffracted by a thin crystal placed in the object plane of the objective lens (Fig. 7). The lens forms in its image plane a magnified and inverted image $\psi(\vec{r}')$ of the wavefunction $\psi(\vec{r})$ emerging from the crystal. In the back focal plane the electron beams converge and form a Fraunhofer diffraction pattern (Fig. 6), representing a Fourier transform $\tilde{\psi}(\vec{q})$ of the wave $\psi(\vec{r})$. The path from the focal

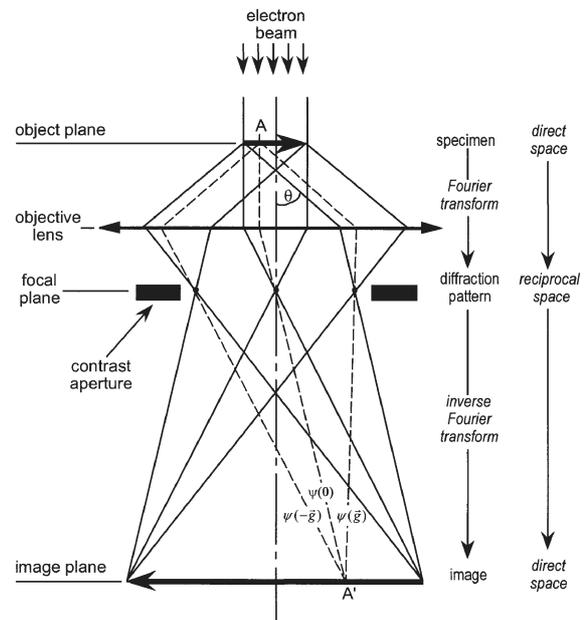


Fig. 7. Image forming by a lens without aberrations



plane to the image plane may be described as an inverse Fourier transform. Thus, for the overall relation between the object and the image we obtain:

$$\psi(\vec{r}') = \mathbf{F}^{-1} \{ \tilde{\psi}(\vec{q}) \} = \mathbf{F}^{-1} \{ \mathbf{F} \{ \psi(\vec{r}) \} \} \quad (12)$$

As the illumination of the object is coherent, the image A' of each of the points A results from an interference of several (spherical) waves admitted by the opening of the contrast aperture (Fig. 6b).

The effect of aberrations intervenes at the level of the focal plane through the action of the transfer function of the microscope $T(\vec{q}, \Delta z)$ (11). The relation (16) takes a more precise form:

$$\psi(\vec{r}') = \mathbf{F}^{-1} \{ D(\vec{q}) \cdot T(\vec{q}, \Delta z) \cdot \mathbf{F} \{ \psi(\vec{r}) \} \} \quad (13)$$

Where $D(\vec{q})$ is an aperture function (equal to unity in the opening of the aperture and zero in the rest of the plane), which describes the limitations due to the insertion of the objective aperture [13, 14].

3. 3. Real lenses and their aberrations

The aberrations of real lenses deform the electron waves and decrease the resolution of the microscope. Fortunately, in the transmission electron microscopy the electron beams propagate close to the optical axis and under small angles. Thus we can neglect higher order aberrations, common in the photon optics. Nevertheless we have to take into account these “axial” aberrations: defocus, spherical aberration, chromatic aberration, astigmatism and coma. The last two aberrations – astigmatism and coma - can be corrected.

3. 3. 1. Spherical aberration

The effect of the spherical aberration (defined by a spherical aberration constant C_s) is to draw the marginal electrons (electrons more distant from the optical axis) more to the optical axis than other ones. Consequently the phase shift of the beams propagating under a different angle θ with respect to the optical axis is given [12] by:

$$\chi_1(\theta) = -\frac{2\pi}{\lambda} C_s \frac{\theta^4}{4} \quad (14)$$

3. 3. 2. Defocus

The focal distance of the electron lenses depends on the excitation current. The variation of the focal distance results in another phase shift

$$\chi_2(\theta, \Delta z) = -\frac{2\pi}{\lambda} \Delta z \frac{\theta^2}{2} \quad (15)$$

which can in certain manner compensate the phase shift due to the spherical aberration, if Δz is negative (lower current = weaker lens).

3. 3. 3. Coherent transfer function

If we consider illumination conditions as perfectly coherent, the incident beam is:

- *monochromatic* (temporal coherence) and
- *parallel* (spatial coherence)

there are both phase shifts $\chi_1(\theta, \Delta z)$ (14) and $\chi_2(\theta, \Delta z)$ (15) in the argument of the coherent transfer function T_c :

$$T_c(\theta, \Delta z) = \exp i(\chi_1 + \chi_2) = \exp \left\{ \frac{2\pi}{\lambda} \left(C_s \frac{\theta^4}{4} + \Delta z \frac{\theta^2}{2} \right) \right\} \quad (16)$$

Scherzer [12] has shown that there is a particular defocus (Scherzer defocus)

$$\Delta z = -1.2 \sqrt{C_s \lambda} \quad (17)$$

for which the microscope transfers a wide band of the spatial frequencies. The imaginary part of the transfer function is plotted in Fig. 8. The first zero is considered to be the limit of the point resolution of the microscope. However it is possible, taking into account some precautions, to use also spatial frequencies higher than this Scherzer limit.

3. 4. Partially coherent illumination

The real incident electron beam is only partially coherent. The effect of the temporary coherence is linked with the dispersion of electron energy and the chromatic aberration of the objective lens. The spatial coherence is the imperfection in the parallelism of the incident beam.

3. 4. 1. Chromatic aberration

The focal length of an electron lens depends on the energy of electrons. The total dispersion of the energy of electrons is given by a combination of several contributions: the initial dispersion ΔE when the electrons leave the cathode, variations of the accelerating voltage ΔV , fluctuation of the lens current ΔI and the energy loss resulting from the interactions of the electron with the specimen.

This energy dispersion causes a difference in the focal length Δf , and therefore a point in the object plane is imaged as a disc of the radius Δr . If we neglect the energy loss in the observed sample, in the first approximation we can write:

$$\Delta f = C_c \left[\left(\frac{\Delta V}{V} \right)^2 + \left(\frac{2\Delta I}{I} \right)^2 + \left(\frac{\Delta E}{E} \right)^2 \right]^{1/2} \quad (18)$$

where C_c is the coefficient of chromatic aberration. Fejes [15] has shown that a Gaussian distribution of this enlargement of the focal length (18) gives an envelope function of the temporal coherence:

$$E_T = \exp \left\{ -\frac{\pi^2}{\lambda^2} (\Delta f)^2 \theta^4 \right\} \quad (19)$$

(Fig. 8b). This envelope function is damping the transfer of higher spatial frequencies and limiting the resolution of the microscope.

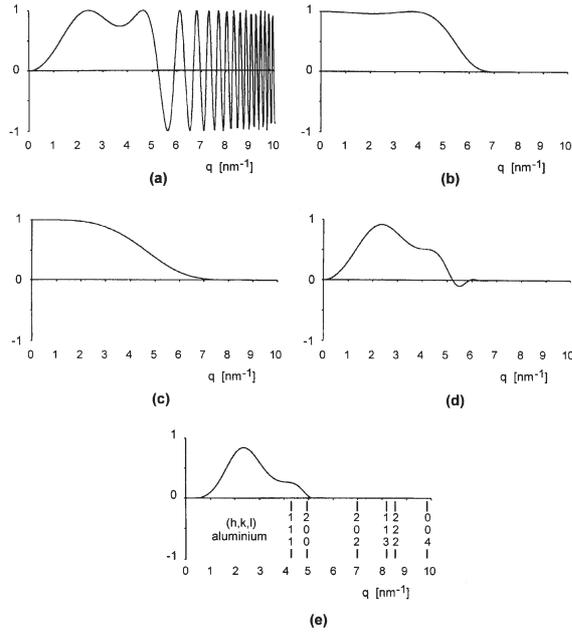


Fig. 8. Contrast transfer function and its envelopes (Scherzer defocus) for a 200 kV microscope with LaB₆ gun, $C_s = 0.5$ mm. (a) contrast transfer function, (b) spatial coherence envelope, (c) temporal coherence envelope, (d) product (a)*(b)*(c) – amplitude transfer function, (e) intensity transfer function with the positions of spatial frequencies corresponding to hkl planes of aluminium.

3. 4. 2. Convergence of the incident beam

The convergence of the incident beam (spatial coherence) limits the resolution of the microscope in a similar manner as the chromatic aberration does [16]. The envelope function of the spatial coherence α for a semi-angle of convergence is given by the expression:

$$E_s = \frac{2J_1(x)}{x} \quad (20)$$

where $J_1(x)$ is the Bessel function of the first order and the first kind, and x is given by :

$$x = 2\pi\alpha \left\{ \Delta z \frac{\theta}{\lambda} + \left[\lambda C_s - i\pi(\Delta f)^2 \right] \frac{\theta^3}{\lambda^2} \right\} \quad (21)$$

The spatial coherence envelope is plotted in Fig. 8c.

3. 4. 3. Real transfer function

Real transfer function is given by the product of the expressions (16) (19) and (20):

$$T_R = T_C \cdot E_T \cdot E_S \quad (22)$$

Its imaginary part is plotted in Fig. 8d. The parameters: $C_s = 0.5$ mm, $\Delta f = 10$ nm and $\alpha = 1$ mrad correspond to the microscope Philips CM20 UltraTwin with a thermoemission LaB₆ gun.

3. 4. 4. Amplitude and intensity transfer

Let us consider a wavefunction $\psi_E(\vec{r})$ at the exit of the crystal and its Fourier transform $\tilde{\psi}_E(\vec{q})$. The influence of the microscope is given by the transfer function:

$$\tilde{\psi}_E(\vec{q})T(\vec{q}, \Delta z).$$

In order to obtain the amplitude of the image $\psi_{IM}(\vec{r}')$, it is necessary to perform an inverse Fourier transform:

$$\psi_{IM}(\vec{r}') = \mathbf{F}^{-1} \{ \tilde{\psi}_E(\vec{q}) \cdot T(\vec{q}, \Delta z) \}$$

Owing to the fact that Fourier transform of a product is a convolution of Fourier transforms of individual functions, it is possible to write (using the relation

$$\mathbf{F}^{-1} \{ T(\vec{q}, \Delta z) \} = h(\vec{r}')$$

$$\psi_{IM}(\vec{r}') = \psi_E(-\vec{r}') * h(\vec{r}') \quad (23)$$

It is the same linear relation as the expression (9), and therefore the *amplitude transfer by the microscope is linear*. However, since the quantity that we detect on the photographic film or on a camera is intensity, we have to study the intensity transfer.

The intensity of the wave emerging from the crystal is $I_E(\vec{r}) = |\psi_E(\vec{r})|^2$, the intensity of its image $I_{IM}(\vec{r}')$ can be expressed with the aid of the relation (23):

$$I_{IM}(\vec{r}') = |\psi_{IM}(\vec{r}')|^2 = |\psi_E(-\vec{r}') * h(\vec{r}')|^2$$

Since the functions $\psi_E(\vec{r})$ and $h(\vec{r}')$ are generally complex, the relation between the intensity of the object $I_E(\vec{r})$ and the intensity of the image $I_{IM}(\vec{r}')$ is not simple.

The intensity transfer under the coherent illumination is not linear. The relation between the object and its image can be established only by the calculation.

However, in special cases the non-linearity of the transfer in intensity is overbalanced by the nature of the object. This is true for isolated atoms or "weak phase objects" - objects of several nanometers thick only [17]. For a detailed analysis of these special cases see [13, 17-19].

If the imaging is incoherent, we obtain an expression $I_{IM}(\vec{r}') = I_E(\vec{r}) * |h(\vec{r}')|^2$ which is similar to (23). Therefore the intensity transfer in this case is linear and thus the Z-contrast imaging gives directly interpretable images (see § 8).

4. Interpretation of the interference image

The interpretation of interference images with an atomic resolution consists in the exact determination of the positions of atomic columns with respect to the black and white contrast on the micrograph. A direct interpretation is possible only in very special cases of "weak-phase objects", where the final image corresponds to the projected potential of the whole specimen.

Real crystals are "strong-phase objects". Neither their interaction with the incident electron beam, nor the transfer of the intensity of the image are linear. It is not possible to



assign intuitively a crystal structure that diffracted the electrons to an experimental image.

The main problem is that the positions of maxima and minima of interference with respect to atomic columns depend at the same time on the crystal thickness and on the defocus of the objective lens. Furthermore, it is often very difficult to determine the values of both these experimental parameters. Fig. 9 shows distribution of minima and maxima of interference with respect to the positions of atomic columns for a wave emerging from a perfect crystal. Independently on the transfer of the contrast by the microscope, the maxima are placed at the position of atomic columns (z_1) or at the "tunnels" between them (z_3). However there are certain thicknesses (z_2) for which the emerging wave has an intermediate form and another series of "supplementary" maxima appears. These three forms of the waves repeat periodically with the crystal thickness.

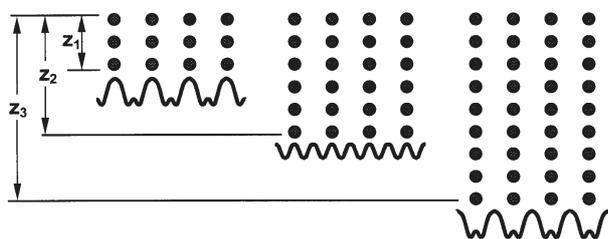


Fig. 9. Maxima and minima of the electron wave emerging from the crystal

Let us consider an electron wave (wavefunction) emerging from the crystal and this wavefunction having its maxima at the position of the atomic columns (Fig. 10). The objective lens creates in its image plane an inverted and magnified image of this wavefunction. The particularity of the interference (coherent) imaging is that we obtain a series of images (called Fourier images) spread around the image plane of the objective lens - Fig. 10. These images are all "sharp". This is an important difference in comparison to the incoherent imaging, giving only one sharp image, under the conditions of exact focus (Gauss focus). Fourier images repeat with a period Δz which depends on the lateral periodicity of the object, represented here by the interplanar spacing d_{hkl} and on the wavelength λ of the electrons:

$$\Delta z = \frac{2d_{hkl}^2}{\lambda}$$

For imaging of $\{111\}$ crystal planes of aluminium with 200 keV electrons we obtain $\Delta z = 43$ nm.

From Figure 10 it is obvious, that depending on the focus of the objective lens we obtain so called "structure images" on which the atoms appear as "white" or "black" dots. Unfortunately there is also a third possibility that the maxima of intensity at the positions of atoms and at the position of "tunnels" between them have approximately the same intensity. In this case the interference image does not correspond to the geometry of the crystal.

For the interpretation of the atomic resolution micrographs it is therefore necessary to create a model of the corresponding crystal and to simulate first its interaction with the electrons and to obtain the wavefunction emerging

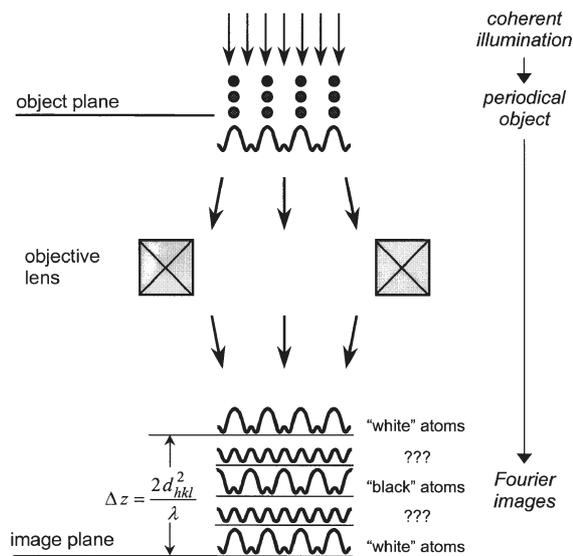


Fig. 10. Fourier images

from the crystal. The next step is to calculate the image of this wavefunction and to compare the result with the experimental micrograph. If the simulation does not correspond to the experimental image, it is necessary to change the model and to repeat the simulation until a reasonable agreement is reached.

This method is rather cumbersome, in particular if we do not know the exact values of certain experimental parameters, like defocus or specimen thickness. In this case it is necessary to calculate several series of simulations for different ranges of parameters.

5. Practical aspects of HREM

In order to obtain good quality interference images, several conditions should be fulfilled. Firstly it is necessary to have a suitable, well aligned, electron microscope. Secondly the sample should be as thin as possible and tilted to a crystal zone axis to align the atomic columns parallel to the electron beam.

Let us comment the choice of the microscope. As the resolution of a TEM depends on the electron wavelength (acceleration voltage) and the quality of the objective lens (C_s):

$$\delta \approx \lambda^{3/4} C_s^{1/4},$$

we have two possibilities: either to use the lowest λ (highest acceleration voltage), or a very low C_s . However, high-voltage TEMs are not always available, and furthermore there is a problem of higher specimen damage due to electron irradiation. For a medium-voltage TEMs (200-300 kV), the lowest commercially available values of C_s are ~ 0.5 mm. (Nevertheless, a special corrector allowing to decrease the C_s value down to zero has been recently developed [20]). The lowest C_s is the best, allowing largest passband at Scherzer defocus (17). On the other hand a very low C_s means also low angle tilting possibilities of the specimen (typical values are about $\pm 15^\circ$ at maximum).

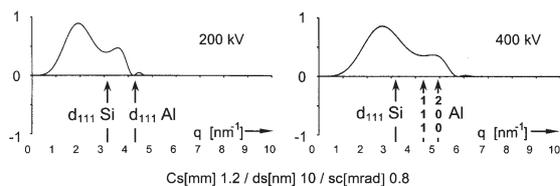


Fig. 11. Contrast transfer function for a 200 kV and a 400 kV microscope with $C_s = 1.2$ mm. Positions of hkl planes of aluminium and silicon are indicated

The choice of the microscope depends also on the interplanar spacing of the studied specimen. If we want to image crystals having large lattice parameters, a higher C_s could be sufficient. Fig. 11 shows a CTF of a 200 kV and a 400 kV TEM with an objective lens having $C_s = 1.2$ mm. It is clearly seen that a 200 kV microscope with this objective lens is not sufficient to image atomic columns of aluminium, however it is sufficient for silicon, having much larger interplanar spacing.

Electron guns

Another factor, which influences the contrast in an important manner, is the electron gun. In general, two kinds of electron guns are available - a thermoemission one (with tungsten or LaB₆ emitter) and a field-emission gun. Table 1 compares parameters of these guns. In HREM, thermoemission guns with a LaB₆ emitter are used, because LaB₆ has a much lower work function than tungsten and therefore the density of electrons is higher, allowing to reduce exposure times.

Table 1. Parameters of electron sources, operating at 100 k

	W	LaB ₆	FEG
Temperature [K]	2700	1700	300
Energy spread [eV]	3	1.5	0.3
Size [μm]	50	10	<0.01
Brightness [$\text{Am}^{-2}\text{srad}^{-1}$]	10^9	5×10^{10}	10^{13}
Current density [Am^{-2}]	5×10^4	10^6	10^{10}

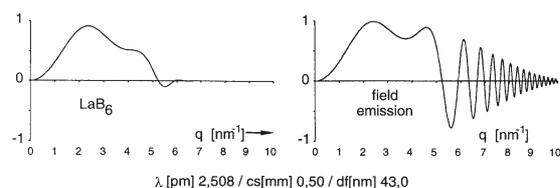


Fig. 12. Comparison of transfer functions of the same microscope fitted with LaB₆ and field emission gun (FEG)

A very expensive field-emission gun (FEG) [4] differs from the thermoemission one by a much smaller size of the source providing much higher spatial coherence (parallelism) of the beam. Also its temporal coherence (energy spread) is better. In consequence, the transfer function of a field-emission gun TEM (FEG TEM) shows many oscillations beyond its first zero (Scherzer limit) - Fig. 12. The microscope is able to transfer information with higher spatial frequencies (lower distances in the direct space). However, due to oscillations in the CTF, this information could be difficult to interpret. This higher information limit is appreciated mainly in electron holography, allowing to "deblur" the wave transfer in the microscope (see § 7).

6. Illustrations of HREM

A disadvantage of HREM is that the image is a 2-dimensional projection. It means that crystal defects should be edge-on and, if possible, extend from the top to the bottom of the specimen. If there are variations in the z direction, the interpretation becomes more difficult. The simulations take also much more time, because it is necessary to calculate the projection of electrostatic potential for many different slices.

Fig. 13 shows a cross-section of a Si/SiO₂ interface prepared by a plasma torch. The interface is edge-on, showing several ledges and a crystallite formed in the amorphous SiO₂. The Si crystal lattice is imaged in the <110> zone axis, each white dot in the micrograph representing two atomic columns (dumbbells) in the diamond structure. In order to separate these columns in the image, a microscope with resolution close to 0.1 nm is required.

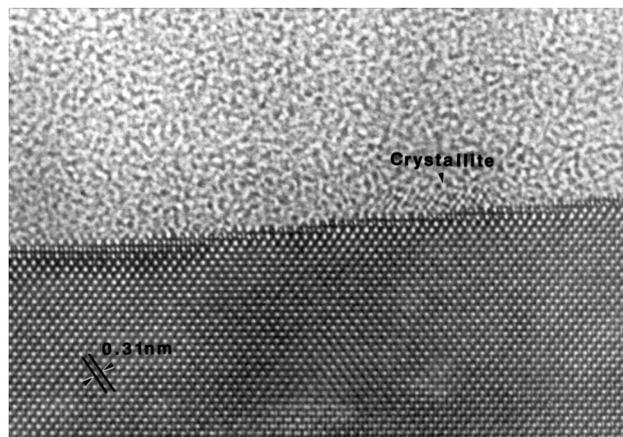


Fig. 13. Crystallite at the interface Si/SiO₂ (courtesy M. Froment, Université Pierre et Marie Curie, Paris)

A next nearest-neighbour antiphase boundary in the ordered alloy based on Fe₃Al is presented in Fig. 14. The crystal defect is edge-on, only Al sublattice of the D0₃ structure is imaged. The arrows indicate parts of the edge-on imaged defect. The inset in the lower part of the micrograph shows the simulation of the atomic configuration of the defect. A defocus-thickness map for this alloy is in Fig. 15. It is clearly seen, that only specific combinations of crystal thickness and defocus of the objective lens give structure images. For more details see [21, 22].

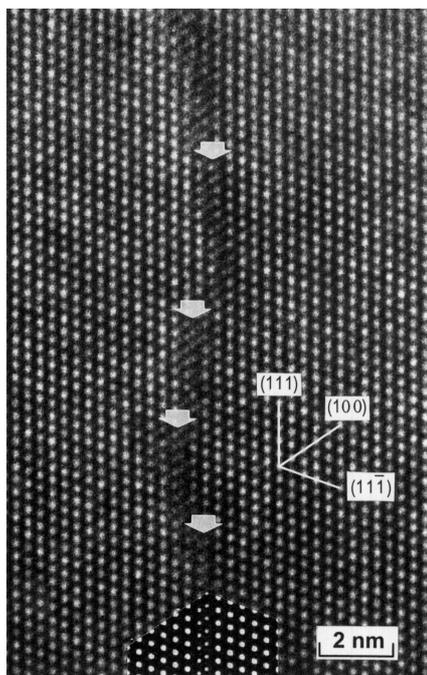


Fig. 14. Next nearest neighbour antiphase boundary in the Fe_3Al based intermetallics. The arrows indicate parts of the defect imaged edge-on. The inset in the lower part of the micrograph shows simulation of the atomic configuration of the defect

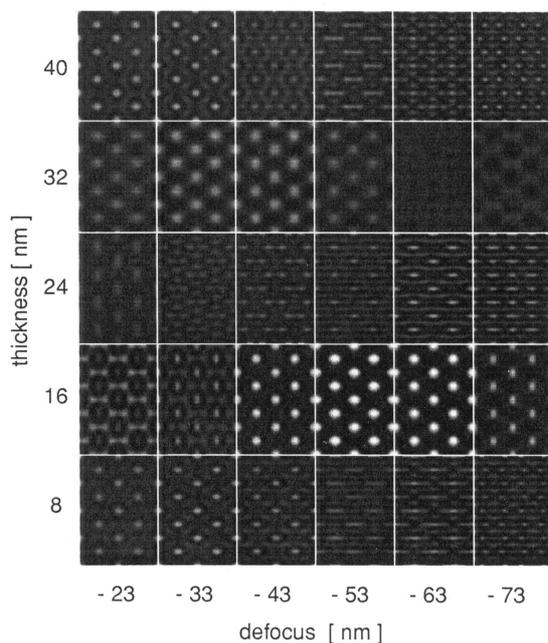


Fig. 15. A defocus - thickness map of the Fe_3Al intermetallics

Fig. 16 illustrates a strong sensitivity of the contrast of plate-like Guinier-Preston (GP) zones in an Al-1.84 at.% Cu alloy on the defocus. The micrograph is interpreted by simulations with models having different concentration of copper in GP platelets. The simulations show a weak sensitivity of the contrast to the copper content in the GP platelets. The best fit is obtained for a "classical" model of 100 % Cu in GP zones. Guinier-Preston zones do not extend

from the top to the bottom of the thin foil and their contrast is very complicated [23, 24].

A Guinier-Preston zone in an Al-1.84 at.% Cu alloy sheared by a dislocation is shown in Fig. 17. The inset shows the configuration before and after the interaction.

7. Electron holography

Although electron holography has been developed in the early 1990s, Gabor [25] had originally proposed this technique in 1948 as a way to improve the resolution of the TEM. The delay in its implementation was due to the lack of sufficiently coherent electron sources - field-emission guns.

Electron holography is based on wave optics. As we have seen in the description of the bright-field phase contrast imaging, the image recorded on photographic plate or CCD-camera is intensity (i.e. the square of the amplitude) of the electron wave. The more essential part of the electron wave - the phase information - is lost. Holographic recording techniques overcome this problem by allowing subsequent reconstruction of the amplitude and phase of the electron wave by computer processing. There are several different methods of electron holography [26]. Principles of one of the most widely used techniques - off-axis electron holography [27, 28] - are outlined in this paragraph.

In order to record the amplitude and the phase of the electron wave, the aberrated image wave is superimposed with an unscattered plane reference wave resulting in an interference pattern - an off-axis electron hologram. In practice, off-axis electron holography is carried out using a field-emission gun TEM (FEG TEM) equipped with a beam splitter, placed below the objective lens. The beam splitter is made by a metal coated glass fiber $< 0.5 \mu\text{m}$ in diameter, assembled to give electron biprism (Fig. 18). The voltage applied on the fiber is positive and ranges from 10 to 150 V. It is important to have the possibility to rotate the biprism or the specimen, in order to align the biprism with the edge of the specimen. A part of the beam Ψ_G passes through the specimen while the other one acts as a reference beam Ψ_R . The amplitude and the phase information are encoded in the contrast modulation and in the bending of the interference fringes between image and reference waves.

The interpretation of an electron hologram consists in transferring the micrograph to a computer for numerical wave-optic processing. This allows to correct the aberrations and to reconstruct the complex object exit wave. After this "deblurring" of the wave transfer in the microscope, final structural fitting, as in the case of conventional bright-field phase contrast imaging (i.e. electron-specimen interaction simulation), is carried out. The advantage is that the number of parameters for the simulations is reduced and quantitative fitting procedures become feasible. Another advantage of electron holography is that it extends lateral resolution beyond the resolution limit of conventional microscopy, achieving resolution close to 0.1 nm using medium-voltage (200-300 kV) microscopes.

Formerly, holography was applicable only by specialists with special instrumentation. Currently, laboratories are equipped with modern instruments which can be easily

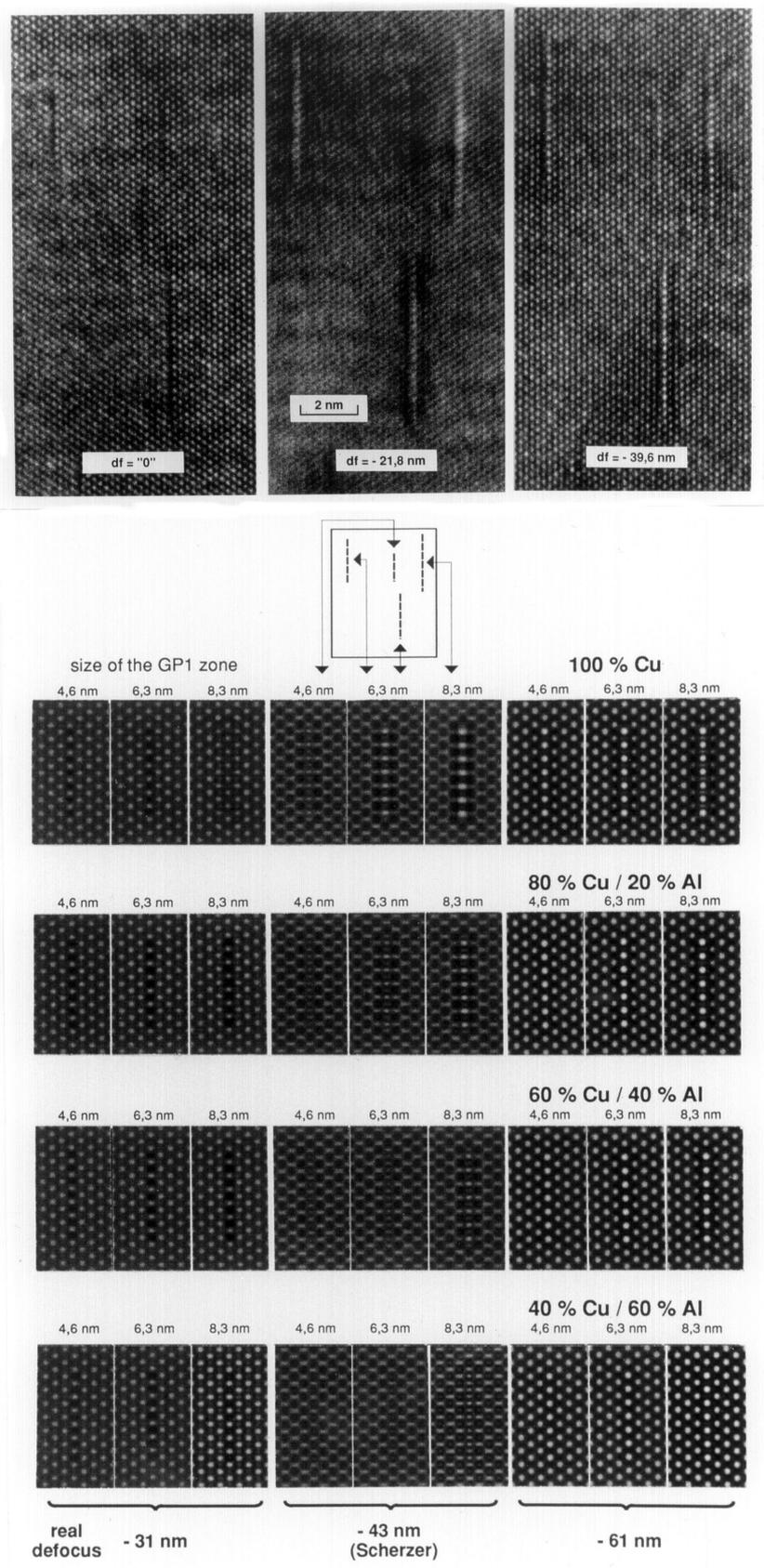


Fig. 16. Plate-like Guinier-Preston (GP) zones in the Al-1.84 at.% Cu alloy. The micrograph is interpreted by simulations with models having different copper concentration in GP platelets

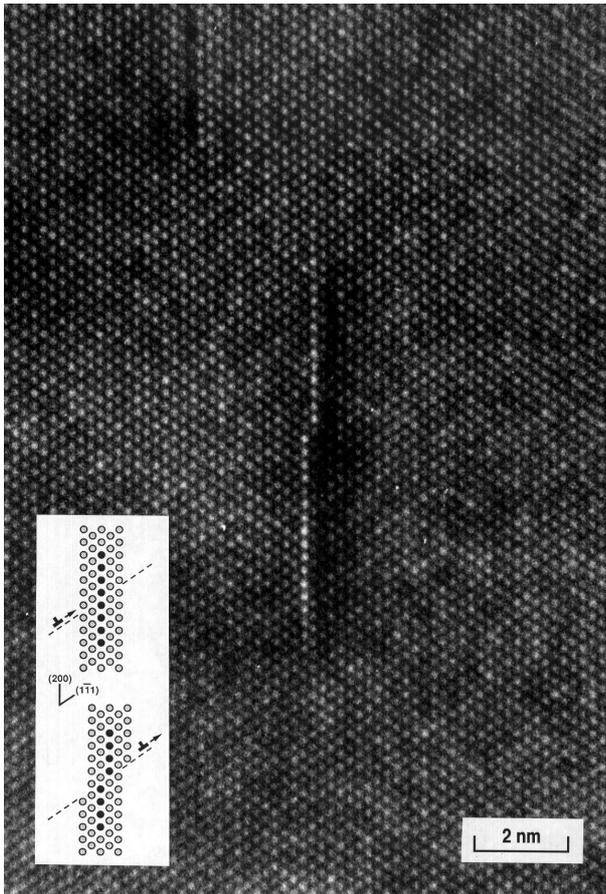


Fig. 17. A Guinier-Preston zone sheared by a dislocation

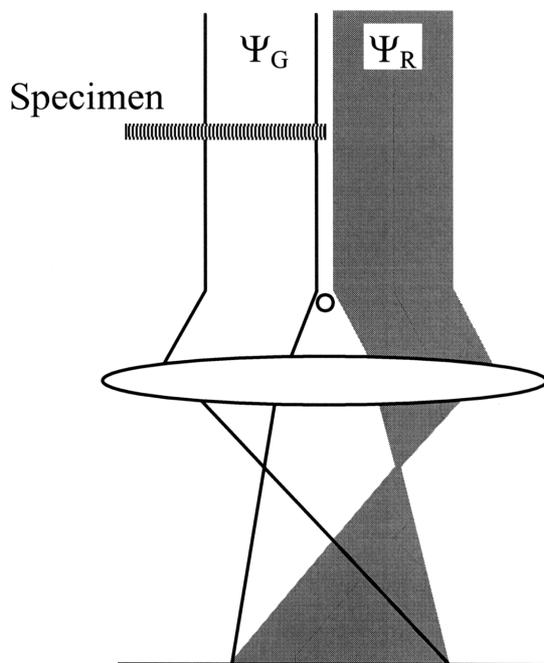


Fig. 18. Principle of off-axis electron holography

adapted for off-axis holography. Basic requirements are a highly-coherent field-emission gun and an electron biprism.

The range of applications is wide. Off-axis electron holography is used for the analysis of structures with atomic resolution. However there are more applications with

lower resolution, such as measurements of doping profiles in pn-junctions, imaging of electric and magnetic fields with lateral resolution of a few nanometers, doping analysis in interfaces etc. [29].

8. Z-contrast imaging

Allowing incoherent imaging, the Z-contrast technique represents a new approach to atomic resolution electron microscopy. It provides directly interpretable images - maps of scattering power of the specimen [30]. This technique is most easily implemented in a (dedicated) scanning transmission electron microscope (STEM) [4]. Many conventional TEMs are also equipped with scanning facilities and electron detectors and they can be used also in a scanning transmission mode [31]. The STEM image is obtained by scanning an electron probe of atomic dimensions across the specimen and collecting electrons by a post-specimen detector (Fig. 19). Scan coils raster the probe across the specimen and their current is used to raster simultaneously the beam of a cathode ray tube (CRT). The signal proportional to intensity of electrons scattered and detected by an electron detector is then displayed on CRT, each pixel on the CRT showing intensity of a part of the image. The persistence of the screen phosphor fluorescence, or frame-store electronics allow the operator to see the whole image. Digital image acquisition is a natural consequence of this mode of operation.

Magnification of the image is given by the ratio of the CRT viewing area dimensions and dimensions of the scanned area. If the scanned area on the specimen is $1 \mu\text{m} \times$

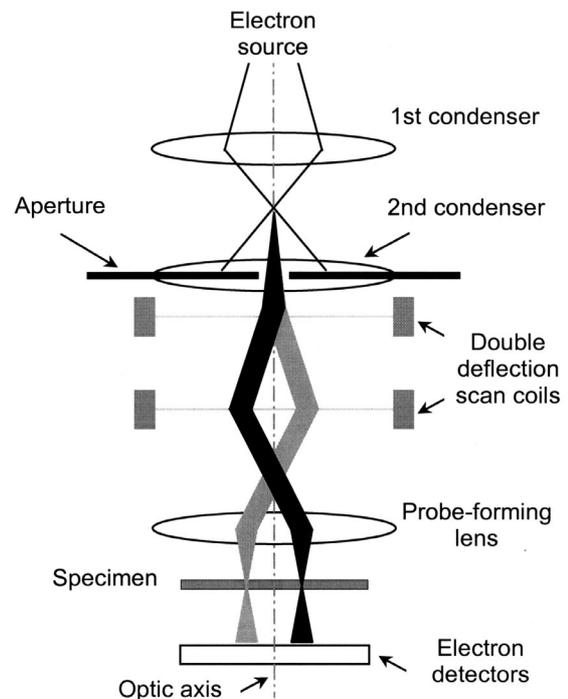


Fig. 19. Principle of a scanning-transmission electron microscope (STEM)

1 μm and the resulting image is displayed on a CRT with the area of 10 cm \times 10 cm, the magnification is 100 000 \times . Resolution of the technique is determined by the size of the focused electron probe.

For incoherent imaging it is necessary to detect incoherent electrons scattered at angles greater than 50 mrad ($\sim 3^\circ$). A special high-angle annular dark field (HAADF) detector was designed to complement standard bright-field (BF) and annular dark-field (ADF) STEM detectors (Fig. 20). In order to image individual atomic columns, the probe needs to be smaller than the columnar spacing of the crystal. The probe has a finite size determined by the geometrical size of the electron source and the magnification settings of the probe-forming lenses. The brightest electron sources are therefore required to yield probes of the smallest dimensions that have sufficient current to form an image. When a field emission gun is used, typical values of an electron probe full-width-half-maximum (FWHM) range from 0.16 to 0.22 nm, with the probe current around 1 nA.

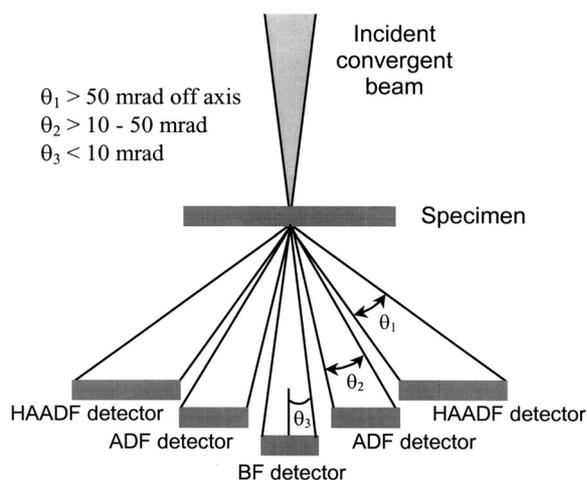


Fig. 20. Configuration of STEM electron detectors

For each probe position, the scattered intensity between the inner and outer angles of the annular detector is summed and displayed on the CRT screen. Each summation, or integration, has the effect of suppressing the interference contrast that tends to produce image artifacts in phase-contrast images. When combined with the channeling effect, this results in an image with an intensity distribution corresponding unambiguously to the location of atomic columns in the structure. Effects of beam damage, contamination and the presence of amorphous surface layers can be minimized.

The method of Z -contrast is attractive whenever the atomic arrangement of crystals, their interfaces and defects is to be ascertained with a degree of chemical sensitivity. Moreover, the detector arrangement allows electrons scattered to low angles to be collected simultaneously by other detectors. This opens the opportunity to do electron energy loss spectroscopy (EELS) and energy-dispersive X-ray spectroscopy (EDS) to gain further compositional data

and/or information on local electronic band structure which may be correlated to the atomic resolution image.

References

- [1] K. Fujiwara, *J. Phys. Soc. Japan* **16** (1961) 2226.
- [2] K. Fujiwara, *J. Phys. Soc. Japan, Suppl. BII*, **17** (1962) 118.
- [3] H.A. Bethe, *Ann. Phys. Lpz.* **87** (1928) 55.
- [4] D.B. Williams & C.B. Carter, *Transmission Electron Microscopy*, Plenum Press, New York, 1996.
- [5] R.J. Glauber, *Phys. Rev.* **91** (1953) 459.
- [6] J.M. Cowley & A.F. Moodie, *Acta Cryst.* **10** (1957) 609.
- [7] K. Ishizuka & N. Uyeda, *Acta Cryst.* **A33** (1977) 740.
- [8] P. Stadelmann, *Ultramicroscopy* **21** (1987), 131.
- [9] M. Born & E. Wolf, *Principles of Optics*, 3rd ed., Pergamon, 1965.
- [10] J.W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1968.
- [11] J.D. Gaskill, *Linear Systems, Fourier Transforms and Optics*, John Wiley, 1978.
- [12] O. Scherzer, *J. Appl. Physics* **20** (1949) 20.
- [13] J.C.H. Spence, *Experimental High-Resolution Electron Microscopy*, Oxford University Press, 1988.
- [14] L. Reimer, *Transmission Electron Microscopy*, Springer-Verlag, 1989.
- [15] P.L. Fejes, *Acta Cryst.* **A33** (1977) 109.
- [16] M.A. O'Keefe & J.V. Sanders, *Acta Cryst.* **A31** (1975) 307.
- [17] K.J. Hanssen & G. Ade, *Optik* **44** (1976) 237.
- [18] J. Deseaux, In: *Microscopie électronique en science des matériaux*, Summer school CNRS, Bombannes, 1981, eds. B. Jouffrey, A. Bourret et C. Colliex, Éditions du CNRS (1983), p. 257.
- [19] P. Buseck, J. Cowley & L. Eyring, *High-Resolution Transmission Electron Microscopy and Associated Techniques*, Oxford University Press, 1988.
- [20] H. Rose, *Optik* **85** (1990) 19.
- [21] M. Karlík & P. Kratochvíl, *Intermetallics* **4** (1996) 613.
- [22] M. Karlík, *Mater. Sci. Eng.* **A234-236** (1997) 212.
- [23] M. Karlík & B. Jouffrey, *Acta Mater.* **45** (1997) 3521.
- [24] M. Karlík, B. Jouffrey & S. Belliot, *Acta Mater.* **46** (1998) 1817.
- [25] D. Gabor, In: *Proc. Roy. Soc. London* **197A** (1949) 454.
- [26] J.M. Cowley, *Ultramicroscopy* **41** (1992) 335.
- [27] H. Lichte & K. Scheerschmidt, *Ultramicroscopy* **47**, (1992) 231.
- [28] H. Lichte, *Ultramicroscopy* **51** (1993) 15.
- [29] H. Lichte & M. Lehmann, In: *Proc. EUREM 12*, Brno, Czech Republic, July 9-14, 2000, Vol. 3, I47.
- [30] S.J. Pennycook & D.E. Jesson, *Acta Metall. Mater.* **40** (1992) S149.
- [31] E.M. James, K. Kishida & N.D. Browning, *JEOL News* **34E** (1999) 16.

