OVERVIEW

Refraction of Electron Rays by Atom and Ray Trajectories in Crystal

By Sigemaro Nagakura*

After a brief description of the multi-slice theory for high resolution electron microscopy, an electron ray theory is developed on the basis of geometrical optics. In this theory the atom is regarded as a thin convex lens and the refraction of electron rays by it is formulated. Then such a concept is extended to crystal in order to describe the ray trajectories of electrons. The result obtained may serve for the intuitive understanding of the propagation of electron wave in crystal.

(Received April 2, 1990)

Keywords: *refraction of electron rays, atom lens, atom chain, trajectory of electron ray in crystal, high resolution electron microscopy, ray theory*

I. Introduction

Recent quality improvements of transmission electron microscopes have made it possible to obtain atomic-scale images of various materials without too much difficulty. Further, programs for the image simulation have been prepared and used to interpret the observed images by comparing the calculated images with the observed. However, the computer-aided image simulation does not always lead to the understanding of physical phenomena taking place in object materials and in electron microscopes. The present article is written in order to assist the understanding of atomic-scale image formation mechanism by depicting the behavior of electrons incident to a crystal.

The theory of image formation in high resolution electron microscopy is divided into two parts. One is the application of classical Abbe's theory of resolution for optical microscope, and takes into account the effect of spherical aberration of objective lens and defocusing. There, Kirchhoff's theorem of the wave propagation in vacuum plays the most important role. The other is to investigate how a vacuum wave incident to the entrance surface of an object reaches the exit surface and how it is modulated before leaving for vacuum to form diffraction pattern or image. This problem is to solve the wave equation in the object. Classically, for crystalline objects dynamical diffraction theory was constructed by Darwin, Ewald and Laue for X-rays, and later it was applied by Bethe to electrons. As for such classical works readers should refer appropriate text books⁽¹⁾. Although the theory is rigorous, its application to the atomic-scale image formation in electron microscope was not always

appropriate. As a more practical way, the multi-slice theory was presented by Cowley and Moodie⁽²⁾ by introducing the concept of projected potential. Nowadays, this theory is widely used for the simulation of structure images.

In spite of its usefulness the multi-slice theory has a weakness that it does not lead to "intuitive understanding" of the behavior of electrons in crystal. This is because the theory treats the wave propagation in crystal. A more intuitive understanding may be obtained by introducing the concept of electron ray based on geometrical optics, in place of the concept of wave, under some degree of approximation.

In what follows, a brief description of multi-slice theory is made at first in order to make clear its relation with the following ray theory, and then the ray theory is described, where the atom is regarded as a lens refracting electrons and as an extension of this concept ray trajectories in crystal are discussed.

II. Wave Equation and Multi-Slice Theory

The wave equation for an electron with charge $-e$ accelerated by a potential $-E$ can be written as

$$\Delta\psi(r) + [2\pi K n(r)]^2 \psi(r) = 0 \quad (1)$$

with

$$n(r) = 1 + \delta(r) = 1 + \frac{V(r)}{2E} \quad (2)$$

Here, $\psi(r)$ is the wave function at a point r , $K = 1/\lambda = 2me/h^2$ is the wave number in vacuum with wavelength λ for the electron with moving mass m , h Planck constant, $n(r)$ the refractive index, $\delta(r)$ the increment of refractive index from unity and given by $V(r)/2E$, where $V(r)$ is the relativistic potential in an object. For $n(r)$ is used the approximation of $V(r)/2E$ being

* Nagaoka University of Technology, Kamitomioka, Nagaoka-shi 940-21, Niigata-ken, Japan.

very small. Equation (1) is nothing but Schrödinger equation in a steady state, but it is rewritten in the form used in wave optics.

In order to solve (1) for a parallel-sided crystal, in the multi-slice theory, the object crystal is cut into many thin slices in parallel with the surface, on which the coordinates (x, y) are taken, while the z -axis is the normal of the surface. Each slice thickness is t as depicted in Fig. 1. Next, the slice is divided along the z -direction into many narrow columns, in each of which the wave function is assumed not to vary in the x - y plane. Then, the wave function applicable to the respective columns can be separated as $\psi(x, y, z) = \psi(x, y) \cdot \psi(z)$, and so (1) becomes

$$\frac{\partial^2 \psi(x, y)}{\partial x^2} + \frac{\partial^2 \psi(x, y)}{\partial y^2} = 0, \quad (3)$$

and

$$\frac{\partial^2 \psi(z)}{\partial z^2} + [2\pi K \{1 + \delta(x, y, z)\}]^2 \psi(z) = 0. \quad (4)$$

A simple form of the solution of (4) can be obtained if we can make $\delta(x, y, z)$ independent of z . For this purpose, it is averaged over t as

$$\langle \delta(x, y, z) \rangle_t = \frac{1}{t} \delta_p(x, y) = \frac{1}{t} \int_0^t \delta(x, y, z) dz, \quad (5)$$

and this is used in place of $\delta(x, y, z)$. We call $\delta_p(x, y)$ the increment of projected refractive index, which is related by $\delta_p(x, y) = V_p(x, y)/2E$ with the projected potential $V_p(x, y) = \int_0^t V(x, y, z) dz$. Then, the solution for the column at (x, y) with height t becomes

$$\psi(x, y, t) = \exp[-2\pi i K t] \exp[-2\pi i K \delta_p(x, y)]. \quad (6)$$

Here, the solution of (3) is taken for respective columns as

$$\psi(x, y) = 1. \quad (7)$$

Equation (6) means that an electron incident to the column is not scattered outside but only subjects to the phase change $2\pi K \delta_p(x, y)$ during passing through the column. Next, each column is compressed in such a way that $t \rightarrow 0$ and $\langle \delta(x, y, z) \rangle_t \rightarrow \infty$, keeping $\delta_p(x, y)$ constant. The

resulting function is called the transmission function and written as

$$q(x, y) = \exp[-2\pi i K \delta_p(x, y)]. \quad (8)$$

Thus, the slice is regarded as a plane on which $\delta_p(x, y)$ is distributed to give the incident electron a phase change.

The wave at a point (x', y', z') at a distance z' from the plane is determined by applying Kirchhoff's theorem. If we take the origin on the plane of $\delta_p(x, y)$ and use the Fresnel approximation, we have

$$\begin{aligned} \psi(x', y', z') &= \exp[-2\pi i K z'] \frac{iK}{z'} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} q(x, y) \\ &\times \exp\left[-\pi i \frac{K}{z'} \{(x' - x)^2 + (y' - y)^2\}\right] dx dy. \end{aligned} \quad (9)$$

This can be calculated by giving $q(x, y)$.

Let the incident slice be the 0-th plane, the next one be the 1-st, and so on. Then, if we put the incident wave as $\psi(x_0, y_0) = 1$, the transmitted wave of the 0-th plane is $q(x_0, y_0) = \exp[-2\pi i \delta_p(x_0, y_0)]$ in accordance with (8). Since the distance from the 0-th plane to the 1-st is t , the incident wave to the 1-st plane from the 0-th can be written as

$$\begin{aligned} \psi(x_1, y_1, t) &= \exp[-2\pi i K t] \frac{iK}{t} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} q(x_0, y_0) \\ &\times \exp\left[-\pi i \frac{K}{t} \{(x_1 - x_0)^2 + (y_1 - y_0)^2\}\right] dx_0 dy_0 \end{aligned} \quad (10)$$

from (9) by putting $z' = t$. The transmitted wave of the 1-st plane is

$$q(x_1, y_1) = \exp[-2\pi i K \delta_p(x_1, y_1)] \psi(x_1, y_1, t). \quad (11)$$

Therefore, the incident wave to the 2-nd plane becomes

$$\begin{aligned} \psi(x_2, y_2, t) &= \exp[-2\pi i K t] \frac{iK}{t} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} q(x_1, y_1) \\ &\times \exp\left[-\pi i \frac{K}{t} \{(x_2 - x_1)^2 + (y_2 - y_1)^2\}\right] dx_1 dy_1. \end{aligned} \quad (12)$$

The transmitted wave of the 2-nd plane is

$$q(x_2, y_2) = \exp[-2\pi i K \delta_p(x_2, y_2)] \psi(x_2, y_2, t). \quad (13)$$

Repeating such a procedure the wave on the exit plane is determined. Finally, this wave leaves the exit surface into vacuum, and at a long distance it produces a Fraunhofer diffraction pattern.

III. Refraction of Electron Rays by Atom

1. General

The multi-slice theory is an ingenious and practical way to calculate the propagation of electron wave through a crystal. However, the concept of wave propa-

Fig. 1 Slice made of one atom layer. Thickness is t . One of the columns is hatched.

gation does not straightforwardly lead to the intuitive understanding of related physical phenomena. This is because, in order to determine the incident wave $\psi(x_i, y_i, t)$ at a point (x_i, y_i) on the i -th plane, we have to know Huygens wavelet $(iK/t)q(x_{i-1}, y_{i-1}) \exp[-\pi i(K/t)\{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2\}]$ starting from all the points (x_{i-1}, y_{i-1}) on the $(i-1)$ -th plane. This makes the relationship between the wave on the $(i-1)$ -th plane and that on the i -th obscure. Introduction of the concept of electron ray may make this relationship more clear and may assist the intuitive understanding of electron motion in crystal.

The concept of ray is applicable, if the refractive index of a medium does not vary severely in the range of wavelength. In case of high energy electrons, say 200 keV, the wavelength λ is 2.5×10^{-3} nm, while the atomic radius is about 0.1 nm. Therefore, it can reasonably be assumed that the refractive index does not vary severely in the wavelength range. Therefore, the concept of ray may be applied and the electron trajectory may be depicted according to geometrical optics.

In a strict sense, the concept of ray is applicable only to zero wavelength. In the case of finite wavelength the ray may have a width, within which plane wave nature is kept. The width may be assumed to be the radius of the first Fresnel zone $\sqrt{\lambda t}$, and for $\lambda = 2.5 \times 10^{-3}$ nm and the atomic distance $t = 0.2 \sim 0.4$ nm this is $(2 \sim 3) \times 10^{-2}$ nm, being smaller than the atomic radius. Based on this idea, the ray trajectories are depicted in what follows. However, it must be noted here that the total width of wave incident to the object is so wide as to cover the object, and so each ray is not allowed to exist independently from others but exists as a whole: each ray represents the result of interference among Huygens wavelets. In Appendix, a mathematical treatment of the propagation of waves in the first Fresnel zone is described.

2. Refraction of electron rays by an atom

In the multi-slice theory, the increment of projected refractive index $\delta_p(x, y)$ is made of the z projection of $\delta(x, y, z)$ between two planes separated by a distance t . Therefore, the slice thickness t is constant but $\delta_p(x, y)$ varies from place to place. However, we can assume that $\delta_p(x, y)$ is the projection of such an object that the increment of refractive index is constant but its thickness changes from place to place, *i.e.* a lens. According to this idea, $\delta_p(x, y)$ can be written in the following form:

$$\delta_p(x, y) = \delta \cdot z(x, y), \quad (14)$$

where $\delta = V(0, 0, 0)/2E$ is a constant representing the increment of refractive index and $z(x, y) = V_p(x, y)/V(0, 0, 0)$ is the thickness varying from place to place.

We consider an electron ray incident to an atom. Taking the origin at the center of the atom, the coordinates are so chosen as the x - y plane to be the plane of projection and the z -axis to be the projection direction. Since $\delta_p(x, y)$ is axially symmetric, we consider only $\delta_p(x, 0)$, and for simplicity the incident ray is in the plane of $y=0$. Further, for brevity, the coordinates $(x, 0)$ are replaced by x . In Fig. 2, the curve represents $z(x)$, and the region

Fig. 2 Refraction of electron ray by atom. For convenience, the atom lens is assumed to have a flat entrance surface and a curved exit surface.

between it and the x axis is the region of the atom with constant δ . The line PA is the incident ray falling on the upper flat surface at the point A with the coordinate x . The incident angle is α and the surface normal is n . At the point A the ray is refracted and reaches the point A' with the coordinate x' on the curved bottom surface, where the surface normal is $n(x')$. The angle between $n(x')$ and n is $\omega(x')$. At the point A' the ray is refracted to become $A'Q$. The angle $\chi(x')$ between the refracted and incident rays is the refraction angle, and the angle $\beta(x')$ between the refracted ray and the z -axis is the deflection angle. It is noted that the angles $\omega(x')$, $\chi(x')$ and $\beta(x')$ changes with x' . The arrowhead attached to the arc indicates the sense of the angle.

Under the assumption that the angles α , $\omega(x')$, $\chi(x')$ and $\beta(x')$ are all small, the application of Snell's law to the boundary surfaces gives the relation

$$\beta(x') = \alpha - \chi(x') \quad (15)$$

with

$$\chi(x') = \delta \omega(x') = \left[-\frac{d\delta_p(x)}{dx} \right]_{x'}, \quad (16)$$

because $\omega(x') = [-dz(x)/dx]_{x'}$. Keeping $\delta_p(x)$ constant, if we make $z(x) \rightarrow 0$, we have $x' \rightarrow x$. Therefore, we can remove the prime in these equations, and $\beta(x')$ and $\chi(x')$ can be calculated by knowing $\delta_p(x)$, *i.e.* $V_p(x)$.

Because of the axial symmetry, $\delta_p(x)$ can be expanded as

$$\delta_p(x) = C_0 + C_1 x^2 + C_2 x^4 + C_3 x^6 + \dots + C_n x^{2n} + \dots \quad (17)$$

Comparing this with the equation for the optical convex lens with refractive index n , thickness t , focal length f and spherical aberration coefficient C_s , we can get the relations: $C_0 = (n-1)t$, $C_1 = -1/(2f)$, $C_2 = -C_s/(2f)^2$. The other C_n 's correspond to the higher order spherical aberration coefficients, which cannot be neglected here. Thus, the atom behaves like a convex lens for incident high energy electrons, and so we may call it atom lens.

Fig. 3 Refraction of electron ray by a chain made of N identical atoms. In the right hand side the entrance and exit point x_i , the incident angle $\alpha(x_i)$, and the deflection angle $\beta(x_i)$ are shown.

have

$$\alpha(x_N) = \beta(x_{N-1}) = \alpha_0 - \sum_{i=0}^{N-1} \chi(x_i), \quad (18)$$

and

$$\begin{aligned} x_N &= x_0 + t \sum_{i=0}^{N-1} \beta(x_i) \\ &= x_0 + t \left\{ N\alpha_0 - \sum_{i=0}^{N-1} (N-i)\chi(x_i) \right\}. \end{aligned} \quad (19)$$

These can be calculated by giving x_0 and α_0 . Since the distance between the points x_{i-1} and x_i is $l_{i,i-1} = t[1 +$

IV. Refraction of Electron Rays in Atom Chain

We consider a chain of N atom lenses aligned along the optic axis z with interval t . This can be regarded as N planes, on every of which $\delta_p(x)$ exists, separated by the distance t , as shown in Fig. 3. The 0-th plane is the entrance plane whose coordinate is represented by x_0 , while the N -th is the exit plane whose coordinate is x_N , where $\delta_p(x_N) = 0$. In the right hand side of the figure, the incident and exit point x_i , the incident angle $\alpha(x_i)$ and the deflection angle $\beta(x_i)$ are shown. From these relations we

$\frac{1}{2} \beta(x_{i-1})^2]$, the optical path between the incident point x_0 and the exit point x_N is given by

$$l_{N,0} = t \left\{ N + \sum_{i=0}^{N-1} \left[\frac{1}{2} \beta(x_i)^2 + \frac{1}{t} \delta_p(x_i) \right] \right\}. \quad (20)$$

This determines the phase of ray at the exit point $2\pi K l_{N,0}$. It is noted that $(1/t)\delta_p(x_i) = \langle \delta_p(x_i) \rangle_t$ is the increment of projected refractive index averaged over the distance t at the point x_i .

V. Ray Trajectories in Crystal

1. General

The crystal is composed of various atom chains arranged periodically. Here, for simplicity, we consider the simplest crystal where chains made of N identical atoms are aligned in the z direction and arranged with a distance a . As shown in Fig. 4, $\delta_p(x)$ for the chain at $x=0$ and $\delta_p(x)$ for the chain at $x=a$ overlaps at the middle point M at $x=a/2$, where $[-d\delta_p(x)/dx]_{a/2} = \chi(a/2) = 0$. Hereafter, we call the plane of $\chi(x)=0$ the wall between the atom chains. When the incident ray is in the plane of

Fig. 4 $\delta_p(x)$ in one slice of crystal. $\delta_p(x)$'s from the atom at $x=0$ and $x=a$ overlap to make the wall M at $x=a/2$, where $\chi(a/2)=0$.

Fig. 5 Incident plane inclined to the entrance plane. $P_x A$ and $P_y A$ are respectively the projections of ray PA to the $x-z$ and $y-z$ planes.

wall, it passes through the crystal without suffering the deflection.

As shown in Fig. 5, the incident plane inclines to the entrance plane, in general. In this case, we consider the projections of incident direction PA to the $x-z$ and $y-z$ planes, $P_x A$ and $P_y A$, and put their angles to the z -axis as $\alpha_x(0)$ and $\alpha_y(0)$, respectively. The coordinates of the incident point A is $x_0=(x_0, y_0)$. In a similar way, we denote the incident angles for the i -th plane as $\alpha_x(x_i, y_i)$ and $\alpha_y(x_i, y_i)$, the deflection angles as $\beta_x(x_i, y_i)$ and $\beta_y(x_i, y_i)$, and the refraction angles as $\chi_x(x_i, y_i)$ and $\chi_y(x_i, y_i)$. Then, as an extension of (18), we have

$$\left. \begin{aligned} \alpha_x(x_N, y_N) &= \beta_x(x_{N-1}, y_{N-1}) = \alpha_x(0) - \sum_{i=0}^{N-1} \chi_x(x_i, y_i), \\ \alpha_y(x_N, y_N) &= \beta_y(x_{N-1}, y_{N-1}) = \alpha_y(0) - \sum_{i=0}^{N-1} \chi_y(x_i, y_i), \end{aligned} \right\} \quad (21)$$

with

$$\left. \begin{aligned} \chi_x(x_i, y_i) &= \left[-\frac{\partial \delta_p(x, y)}{\partial x} \right]_{x_i, y_i}, \\ \chi_y(x_i, y_i) &= \left[-\frac{\partial \delta_p(x, y)}{\partial y} \right]_{x_i, y_i}. \end{aligned} \right\} \quad (22)$$

The coordinates of the exit point $x_N=(x_N, y_N)$ and the optical path $l_{N,0}$ are given, as an extension of (19) and (20), by

$$x_N = x_0 + t \sum_{i=0}^{N-1} \beta_x(x_i, y_i), \quad y_N = y_0 + t \sum_{i=0}^{N-1} \beta_y(x_i, y_i), \quad (23)$$

and

$$l_{N,0} = t \left\{ N + \sum_{i=0}^{N-1} \left[\frac{1}{2} \{ \beta_x(x_i, y_i)^2 + \beta_y(x_i, y_i)^2 \} + \frac{1}{t} \delta_p(x_i, y_i) \right] \right\}. \quad (24)$$

Since $\delta_p(x)$ is axially symmetric near the optic axis, the use of the cylindrical coordinates (ρ, ϕ, z) is convenient in such a region. In this case, for the i -th plane, the incident and exit positions are $x_i=(\rho_i, \phi_i)$, the projection angles of the incident direction to the $\rho-z$ and $\phi-z$ planes are $\alpha_\rho(\rho_i, \phi_i)$ and $\alpha_\phi(\rho_i, \phi_i)$, the corresponding deflection angles are $\beta_\rho(\rho_i, \phi_i)$ and $\beta_\phi(\rho_i, \phi_i)$, and the refraction angles are $\chi_\rho(\rho_i, \phi_i)$ and $\chi_\phi(\rho_i, \phi_i)$. Then, instead of (21) and (22), we have

$$\left. \begin{aligned} \alpha_\rho(\rho_N, \phi_N) &= \beta_\rho(\rho_{N-1}, \phi_{N-1}) = \alpha_\rho(0) - \sum_{i=0}^{N-1} \chi_\rho(\rho_i, \phi_i), \\ \alpha_\phi(\rho_N, \phi_N) &= \beta_\phi(\rho_{N-1}, \phi_{N-1}) = \alpha_\phi(0) - \sum_{i=0}^{N-1} \chi_\phi(\rho_i, \phi_i), \end{aligned} \right\} \quad (25)$$

with

$$\left. \begin{aligned} \chi_\rho(\rho_i, \phi_i) &= \left[-\frac{\partial \delta_p(\rho, \phi)}{\partial \rho} \right]_{\rho_i, \phi_i}, \\ \chi_\phi(\rho_i, \phi_i) &= \left[-\frac{\partial \delta_p(\rho, \phi)}{\partial \phi} \right]_{\rho_i, \phi_i}. \end{aligned} \right\} \quad (26)$$

Here, we put $\alpha_p(0)=\alpha_p(0,0)$ and $\alpha_\phi(0)=\alpha_\phi(0,0)$. If $\delta_p(\rho, \phi)$ is axially symmetric, $\chi_\phi(\rho_i, \phi_i)=0$, and the coordinates of the exit point and the optical path are respectively given by

$$\rho_N = \rho_0 + t \sum_{i=0}^{N-1} \beta_p(\rho_i, \phi_i), \quad \phi_N = \phi_0 + N\alpha_\phi(0), \quad (27)$$

and

$$l_{N,0} = t \left\{ N + \sum_{i=0}^{N-1} \left[\frac{1}{2} \{ (\beta_p(\rho_i, \phi_i))^2 + (\alpha_\phi(0))^2 \} + \frac{1}{t} \delta_p(\rho_i, \phi_i) \right] \right\}. \quad (28)$$

If we use the unit vectors $\alpha(x_i) = (\alpha_x(x_i, y_i), \alpha_y(x_i, y_i), 1)$, $\beta(x_i) = (\beta_x(x_i, y_i), \beta_y(x_i, y_i), 1)$ and $\chi(x_i) = (\chi_x(x_i, y_i), \chi_y(x_i, y_i))$ under the approximation of $\alpha_x(x_i, y_i)$, etc. being all small enough, (21) and (22) can be written as

$$\begin{aligned} \alpha(x_N) &= \beta_x(x_{N-1}) = \alpha_x(0) - \sum_{i=0}^{N-1} \chi(x_i) \\ &= \alpha_x(0) + \sum_{i=0}^{N-1} \text{grad } \delta_p(x_i) \end{aligned} \quad (29)$$

2. Axial incidence

At first, we consider the case where the incident ray is parallel to the optic axis, *i.e.* $\alpha_x(0)=\alpha_y(0)=0$. The wall is assumed to be perpendicular to the x -axis, and the ray trajectory in the plane of $y=0$ is considered. Then, from the symmetry of $\delta_p(x, y)$, (18) reduces to

$$\alpha_x(x_N) = \beta_x(x_{N-1}) = - \sum_{i=0}^{N-1} \chi_x(x_i). \quad (30)$$

As shown in Fig. 6, the ray incident to the point M_0 with $x=(a/2)$ on the wall passes through the crystal without deflection because of $\chi(a/2)=0$, while the ray incident to the point A_0 (A_0') near the wall is deflected inward only slightly because of small $\sum_{i=0}^{N-1} \chi_x(x_i)$, and reaches the point A_N (A_N') on the exit plane. The ray incident to the point B_0 with coordinate x_0 ($B_0': -x_0$), on the other hand, crosses the optic axis at the point B_F on the F -th plane, where $\sum_{i=0}^{F-1} \chi_x(x_i)$ is maximum, and then reaches the point B_G ($-x_G$), ($B_G': x_G$), on the G -th plane, where $\sum_{i=0}^{G-1} \chi_x(x_i)=0$. After that, the ray trajectory becomes opposite to the trajectory $B_F B_G$ ($B_F' B_G'$). The trajectory oscillates with the period $4(G-F)t$ and the amplitude x_G . The point B_F is the first focus of this ray. When the incident point approaches to the optic axis such as C_0 (C_0'), the oscillation period becomes shorter and the amplitude smaller. It must be noticed that the coordinate $|x_N|$ for the exit point B_N (B_N') does not become larger than the corresponding coordinate $|x_G|$, *i.e.* $|x_N| \leq |x_G|$. This means that the atom chain behaves like a convex lens.

3. Oblique incidence: Special cases

(1) The case of the incident plane containing the optic axis

Let the incident plane be the plane of $y=0$ and the incident angle be $\alpha_x(0)$. As shown in Fig. 7, when the inci-

Fig. 6 Schematic illustration of the ray trajectories in an atom chain in crystal. The case of axial incidence. The measure along the z axis is very much compressed compared with the measure along the x axis.

dent point is A_0 (coordinate x_0), at first, the ray proceeds outward and reaches the point $A_G(x_G)$ on the G -th plane, where $\sum_{i=0}^{G-1} \chi_x(x_i)=0$ and $x_G > x_0$. After that, the ray trajectory oscillates with amplitude x_G and with a longer period than the case of axial incidence. On the other hand, when the incident point is A'_0 ($-x_0$), the ray crosses the optic axis and then reaches the point $A_{G'}(x_{G'})$ on the G' -th plane, where $\sum_{i=0}^{G'-1} \chi_x(x_i)=0$. After that the ray trajectory oscillates with amplitude $x_{G'}$ and a shorter period than the case of axial incidence.

If the incident point is M'_0 near the wall, after some travel, the ray proceeds along the wall $M_0 M_N$. If the incident position is between M'_0 and M_0 , the ray enters the adjacent atom chain.

(2) The case of the normal of incident plane at the incident point crossing the optic axis

Let the incident points B_0 with coordinates $(0, y)$ or $(\rho_0, \pi/2)$ and B'_0 with $(0, -y)$ or $(\rho_0, -\pi/2)$ and assume $\delta_p(x, y)$ to be axially symmetric. In this case, in (25), $\alpha_p(0)=0$, and if

$$\alpha_p(\rho_N, \phi_N) = \beta_p(\rho_{N-1}, \phi_{N-1}) - \chi_p(\rho_{N-1}, \phi_{N-1}) = 0, \quad (31)$$

then the ray is not deflected in the radial direction. On the other hand, we have from (27)

$$\phi_N = \pm \frac{\pi}{2} + N\alpha_\phi(0). \quad (32)$$

Therefore, the ray trajectory simply makes a spiral along

Fig. 7 Schematic illustration of the ray trajectories in an atom chain in crystal. Special cases of the oblique incidence. The incident plane for the rays incident to the points M'_0 , A'_0 and A'_0 contains the optic axis. The ray trajectories $A_0 A_N$ and $A'_0 A'_N$ cross the optic axis. The incident planes to the points B_0 and B'_0 are parallel to the $x-z$ plane. The ray trajectories are a spiral with pitch Pt . The measure along the z axis is very much compressed compared with the measure along the x axis.

the optic axis as shown in Fig. 7 by the spirals $B_0 B_N$ and $B'_0 B'_N$. The two trajectories cross at a certain point. Their pitch Pt depends on the coordinate of the incident point ρ_0 , and is determined by the relation $Pt\alpha_\phi(0) = 2\pi\rho_0$.

4. Oblique incidence: general case

In general, the incident plane makes some angle with the entrance plane. In this case, the ray trajectory takes a form determined by a combination of oscillation in the radial direction, rotation around the optic axis and translation along the optic axis. Therefore, it is very complex. In Fig. 8, two examples are depicted. The trajectory for the ray incident to the point A_0 takes a form of severely deformed spiral with a longer pitch and a larger amplitude than the spiral shown in Fig. 7. The deformation is caused by the fact that the ray must cross the optic axis. Similarly, the trajectory for the ray incident to the point A'_0 takes a form of severely deformed spiral with a shorter pitch and a smaller amplitude.

The ray trajectory varies depending upon the incident point. In the case of oblique incidence, therefore, so many ray trajectories of different kinds are formed in the

Fig. 8 Schematic illustration of the ray trajectories in an atom chain in crystal. General case of the oblique incidence. The ray trajectories are deformed spirals crossing the optic axis. The measure along the z axis is very much compressed compared with the measure along the x axis.

atom chain, but the rays proceed, as a whole, along the optic axis, although the rays incident near the wall enter the adjacent chains across the wall.

5. Intensity distribution and diffraction

When the length of atom chain is short enough, the ray trajectories are almost parallel one another, except the rays incident to the region very near the optic axis. Therefore, the distribution of rays on the exit plane is almost equal to that of incident rays, and homogeneous. Each ray is different only in its phase. Such a chain length corresponds to the kinematical region in diffraction theory.

With increasing chain length the ray trajectories tend

to curve toward the optic axis. Such a tendency begins to appear from the rays incident to the region near the optic axis and spreads gradually to those incident to the region near the wall. This phenomenon corresponds to the dynamical diffraction or the multiple scattering phenomenon of wave. When the length of atom chain is long enough, independent of the incidence direction, the density of rays on the exit plane is high around the optic axis and low around the wall, as a result of the fact that the atom chain behaves like a convex lens. The phase of each ray is not related in a simple form with $\delta_p(x, y)$; as seen in (24), the optical path is a function of $\sum_{i=0}^{N-1} \delta_p(x_i, y_i)$. In the case of a thick crystal, therefore, we cannot get any detailed information about the potential distribution inside the atom chain from the measurement of the intensity distribution inside the image of atom chain.

The width of each ray leaving the exit plane widens with distance and overlaps to interfere. At a distance of 10 nm, the width becomes $\sqrt{\lambda}t = 0.16$ nm for 200 keV electrons. The interference produces a Fresnel diffraction pattern with approximate correspondence with the distribution of ray density in each atom chain. At a distance very far from the exit surface, the rays from each atom chain so widen that they overlap to cause the Fraunhofer diffraction. Since the ray distribution is the same for each atom chain in the present crystal and periodic, observed diffraction patterns become spotty.

It is added that when each of the rays incident to the crystal satisfies the condition $a\alpha_x(0) = n\lambda$ (n : integer) on the entrance plane, each of the exit rays satisfies the same condition on the exit plane. This is because, as can be seen from (23) and (21), the respective exit points displace from the corresponding entrance points by the same amount $x_N - x_0 = t \{N\alpha_x(0) - \sum_{i=0}^{N-1} \chi(x_i)\}$.

VI. Discussion and Conclusion

In the present article, the concept of electron ray was introduced, and the refraction by atom and the ray trajectory in crystal have been investigated. The theory is based on the geometrical optics, and since only the first Fresnel zone is considered, the degree of approximation is worse than the wave theory, and the results are rather qualitative.

We assumed the width of wave showing the plane wave nature to be $\Delta x = \sqrt{\lambda}t$. According to the uncertainty principle, this corresponds to the uncertainty of wave number $\Delta k_x = 1/(\sqrt{\lambda}t)$. For $t = 2 \sim 4$ nm and $\lambda = 2.5 \times 10^{-3}$ nm, this is $50 \sim 30$ nm $^{-1}$, covering the range of reciprocal space $-20 \sim 20$ nm $^{-1}$ usually taken into account for the calculation of propagating waves. Accordingly, the direction of ray trajectory in the present theory represents the average direction of waves scattered in various directions.

Regarding the atom as a thin convex lens, the present theory is developed, and the ray trajectories in the atom chain is depicted. A consideration shows that the present problem is almost the same as the problem to investigate the motion of electrons in an axially symmetric potential

field, because $\delta_p(\rho)/t$ corresponds to the cylindrical potential field $V_p(\rho)/t$ and $\chi(\rho)$ is proportional to the electric field $-\text{grad}[V_p(\rho)/t]$. Therefore, the present problem can also be solved along the way of classical mechanics⁽³⁾, and the result represents the motion of center of gravity of electron wave giving the ray trajectory described in this article.

The present theory shows that most of the electron rays incident to a crystal with a small angle can be regarded as to proceed in the respective atom chains separated by the wall determined by the condition $\chi(x) = -\text{grad} \delta_p(x) = 0$; in other words, the respective atom chains play the role of a pipe to carry electrons from the entrance surface to the exit surface. This conclusion is correct not only for the case of crystal made of identical atom chains but also for crystals made of different atom chains. Here, as a special case, we consider the case of cross-sectional observation of two crystals A and B being in an epitaxial relation. At the boundary, $\delta_p(A)$ is modified by $\delta_p(B)$ as a result of overlapping, and if the atom A is lighter than B, the modification may be larger for $\delta_p(A)$ than for $\delta_p(B)$ and the wall may displace towards the atom chain A. Such a modification of δ_p taking place at the boundary may reduce the distinction between the atom chains A and B in some extent and influences the image contrast.

In the above the absorption of electron has not been taken into consideration. It can be taken into account by assuming $\delta_p(x)$ to be complex and its imaginary part relates with absorption. It may be reasonable to assume that the imaginary part is large in the neighborhood of optic axis but small in the neighborhood of the wall. In a thick crystal, therefore, the rays crossing the optic axis many times may be severely absorbed and do not reach the exit surface, but the rays passing through near the wall are absorbed only slightly. This interprets the anomalous transmission of electrons or the electron channeling phenomenon taking place in thick crystals.

In conclusion we can say that the present theory is qualitative but may serve for the intuitive understanding of behavior of the propagation of electron waves in crystal, *i.e.* the motion of electrons in crystal.

Acknowledgment

The author would like to express his thanks to Drs. Y. Hirotsu and T. Ishiguro for their discussion and interest.

REFERENCES

- (1) For example, R. W. James: *Optical Principles of the Diffraction of X-rays*, Bell and Sons, London, (1954); *X-sen Kesshogaku (X-ray Crystallography)*, Ed. by O. Nitta, Maruzen, Tokyo, (1961); P. B. Hirsch, A. Howie, R. B. Nicholson, D. W. Pashley and M. J. Whelan: *Electron Microscopy of Thin Crystals*, Butterworth, London, (1965); L. V. Azaroff, R. Kaplow, N. Kato, R. J. Weiss, A. J. C. Wilson and R. A. Young: *X-ray Diffraction*, McGraw-Hill, New York, (1974).
- (2) J. M. Cowley: *Diffraction Physics*, North-Holland, Amsterdam, (1981).
- (3) H. C. H. Nip, M. J. Hollis and J. C. Kelly: *Phys. Lett.*, **28A** (1968), 324.

Appendix

Propagation of Waves in the First Fresnel Zone

We consider the following wave at a point (x_i, y_i, t_i) on the i -th plane contributed by the waves in the first Fresnel zone around a point $(x_{i-1}, y_{i-1}, t_{i-1})$ on the $(i-1)$ -th plane. The points (x_i, y_i, t_i) and $(x_{i-1}, y_{i-1}, t_{i-1})$ are related by Snell's law.

$$\begin{aligned} \psi(x_i, y_i, t_i) = & \exp[-2\pi i K t] \left(\frac{i}{\lambda t} \right) \int_{x_{i-1}-\frac{\sqrt{\lambda t}}{2}}^{x_{i-1}+\frac{\sqrt{\lambda t}}{2}} \int_{y_{i-1}-\frac{\sqrt{\lambda t}}{2}}^{y_{i-1}+\frac{\sqrt{\lambda t}}{2}} q(x_{i-1}, y_{i-1}, t_{i-1}) \\ & \times \psi(x_{i-1}, y_{i-1}, t_{i-1}) \exp \left[-\pi i \left(\frac{K}{t} \right) \{ (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 \} \right] dx_{i-1} dy_{i-1}. \end{aligned} \quad (\text{A} \cdot 1)$$

Averaging $\psi(x_i, y_i, t_i)$ over the first Fresnel zone for the i -th plane, we have

$$\begin{aligned} \langle \psi(x_i, y_i, t_i) \rangle = & \left(\frac{1}{\lambda t} \right) \int_{x_i-\frac{\sqrt{\lambda t}}{2}}^{x_i+\frac{\sqrt{\lambda t}}{2}} \int_{y_i-\frac{\sqrt{\lambda t}}{2}}^{y_i+\frac{\sqrt{\lambda t}}{2}} \psi(x_i, y_i, t_i) dx_i dy_i \\ = & i \exp[-2\pi i K t] \left(\frac{1}{\lambda t} \right) \int_{x_{i-1}-\frac{\sqrt{\lambda t}}{2}}^{x_{i-1}+\frac{\sqrt{\lambda t}}{2}} \int_{y_{i-1}-\frac{\sqrt{\lambda t}}{2}}^{y_{i-1}+\frac{\sqrt{\lambda t}}{2}} q(x_{i-1}, y_{i-1}, t_{i-1}) \psi(x_{i-1}, y_{i-1}, t_{i-1}) dx_{i-1} dy_{i-1} \left(\frac{1}{\lambda t} \right) \\ & \times \int_{x_i-\frac{\sqrt{\lambda t}}{2}}^{x_i+\frac{\sqrt{\lambda t}}{2}} \int_{y_i-\frac{\sqrt{\lambda t}}{2}}^{y_i+\frac{\sqrt{\lambda t}}{2}} \exp \left[-\pi i K t \left\{ \left(\frac{x_i - x_{i-1}}{t} \right)^2 + \left(\frac{y_i - y_{i-1}}{t} \right)^2 \right\} \right] dx_i dy_i. \end{aligned} \quad (\text{A} \cdot 2)$$

Since $|x_i - x_{i-1}| \ll t$ and $|y_i - y_{i-1}| \ll t$, we can replace $((x_i - x_{i-1})/t)$ and $((y_i - y_{i-1})/t)$ by their average incident angles at the point (x_i, y_i, t_i) on the i -th planes:

$$\begin{aligned} \left(\frac{x_i - x_{i-1}}{t} \right) &= \langle \alpha_x(x_i, y_i, t_i) \rangle, \\ \left(\frac{y_i - y_{i-1}}{t} \right) &= \langle \alpha_y(x_i, y_i, t_i) \rangle. \end{aligned} \quad (\text{A} \cdot 3)$$

Then, we have

$$\begin{aligned} \langle \psi(x_i, y_i, t_i) \rangle = & i \exp[-2\pi i K t] \exp[-\pi i K t \{ \langle \alpha_x(x_i, y_i, t_i) \rangle^2 + \langle \alpha_y(x_i, y_i, t_i) \rangle^2 \}] \\ & \times \langle q(x_{i-1}, y_{i-1}, t_{i-1}) \psi(x_{i-1}, y_{i-1}, t_{i-1}) \rangle. \end{aligned} \quad (\text{A} \cdot 4)$$

Since $\delta_p(x_{i-1}, y_{i-1}, t_{i-1})$ is almost constant in the first Fresnel zone around the point $(x_{i-1}, y_{i-1}, t_{i-1})$, it can be replaced by its average value. Then, we can put as

$$q(x_{i-1}, y_{i-1}, t_{i-1}) = \exp[-2\pi i K \langle \delta_p(x_{i-1}, y_{i-1}, t_{i-1}) \rangle]. \quad (\text{A} \cdot 5)$$

Therefore,

$$\begin{aligned} \langle q(x_{i-1}, y_{i-1}, t_{i-1}) \psi(x_{i-1}, y_{i-1}, t_{i-1}) \rangle = & \left(\frac{1}{\lambda t} \right) \int_{x_{i-1}-\frac{\sqrt{\lambda t}}{2}}^{x_{i-1}+\frac{\sqrt{\lambda t}}{2}} \int_{y_{i-1}-\frac{\sqrt{\lambda t}}{2}}^{y_{i-1}+\frac{\sqrt{\lambda t}}{2}} q(x_{i-1}, y_{i-1}, t_{i-1}) \psi(x_{i-1}, y_{i-1}, t_{i-1}) dx_{i-1} dy_{i-1} \\ = & \exp[-2\pi i K \langle \delta_p(x_{i-1}, y_{i-1}, t_{i-1}) \rangle] \langle \psi(x_{i-1}, y_{i-1}, t_{i-1}) \rangle. \end{aligned} \quad (\text{A} \cdot 6)$$

Since the average incident angle to (x_i, y_i, t_i) is equal to the average deflection angle at $(x_{i-1}, y_{i-1}, t_{i-1})$, we replace $\alpha(x_i, y_i, t_i)$ by $\beta(x_{i-1}, y_{i-1}, t_{i-1})$. Then, (A·4) becomes

$$\begin{aligned} \langle \psi(x_i, y_i, t_i) \rangle = & i \exp \left[-2\pi i K t \left\{ 1 + \frac{1}{2} (\langle \beta_x(x_{i-1}, y_{i-1}, t_{i-1}) \rangle^2 \right. \right. \\ & \left. \left. + \langle \beta_y(x_{i-1}, y_{i-1}, t_{i-1}) \rangle^2) + \left(\frac{1}{t} \right) \langle \delta_p(x_{i-1}, y_{i-1}, t_{i-1}) \rangle \right\} \right] \langle \psi(x_{i-1}, y_{i-1}, t_{i-1}) \rangle. \end{aligned} \quad (\text{A} \cdot 7)$$

This gives the relationship between $\langle \psi(x_{i-1}, y_{i-1}, t_{i-1}) \rangle$ and $\langle \psi(x_i, y_i, t_i) \rangle$. Putting $i=N$, we have for the wave on the exit surface

$$\begin{aligned} \langle \psi(x_N, y_N, t_N) \rangle = i^N \exp \left[-2\pi i K t \left\{ N + \sum_{i=0}^{N-1} \left[\frac{1}{2} (\langle \beta_x(x_i, y_i, t_i) \rangle^2 \right. \right. \right. \\ \left. \left. \left. + \langle \beta_y(x_i, y_i, t_i) \rangle^2) + \left(\frac{1}{t} \right) \langle \delta_p(x_i, y_i, t_i) \rangle \right] \right\} \right] \langle \psi(x_0, y_0, t_0) \rangle. \end{aligned} \quad (\text{A} \cdot 8)$$

$\psi(x_0, y_0, t_0)$ is the incident wave. The phase factor coincides with that given by (24). $i^N = \exp [N\pi i / 2]$ is common for all the rays, and no influences are introduced to the result.