

AD 717697



AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
L. G. HANSCOM FIELD, BEDFORD, MASSACHUSETTS

On the Foundations of Crystal Optics
Part I—Dispersion Theory
Part II— Theory of Reflection and Refraction

P. P. EWALD

Part I - Dispersion Theory Translated by Dr. Lowell Hollingsworth from
Annalen der Physik, 49, 1916, No. 1

Part II - Theory of Reflection and Refraction Translated by Dr. Lowell Hollingsworth
from Annalen der Physik, 49, 1916, No. 2

Postscript 1970 by P.P. Ewald

This document has been approved for public
release and sale; its distribution is unlimited.

United States Air Force



Abstract 1970

This two part paper investigates the propagation of light in the visible region through a crystalline medium.

In Part I, which is a slightly shortened version of the author's doctoral thesis of 1912, the objective is to determine if the anisotropic arrangement of ordinary (isotropic) dipoles at the nodal points of an orthorhombic lattice would account for the existence of double refraction. The value of the computed birefringence is found to be comparable to the observed value. Additionally some features of the traditional "theory of dispersion" are disclosed and clarified. In the older theories the "incident optical wave" plays a role but it is shown that in a medium which extends to infinity in all directions no such wave should be assumed: The refractivity is shown to arise as an internal property of the medium.

In Part II, a crystalline medium is considered as filling a half-space and having a plane boundary at $z = 0$. A plane optical wave is incident on this medium. Because of the linearity of the equations it has to be superimposed on the field originating in the crystal. It is shown that this incident optical wave is actually prevented from entering the crystal because of the modification produced in the field of the crystal by the introduction of a boundary. Boundary waves are found to exist on both sides of the boundary. The higher their order, the more rapidly they attenuate as a function of distance away from the boundary but the zero order waves are ordinary undamped plane waves of vacuum velocity c . The fields outside and inside

Abstract 1970

the medium: are connected by the Fresnel formulae: These follow from the condition that the optical field inside the medium and the oscillations generated there form a self-consistent system.

Material Added 1970. The conclusion that the incident optical field cannot penetrate the crystal boundary, together with a similar conclusion in a paper by Oseen, is the basis of the Ewald-Oseen Extinction Theorem.

The same methods and field transformations developed in this two part paper were applied in two later papers which were published as Parts III and IV. These extend the treatment to the passage of X-rays through crystals. "Postscript 1970" published here indicates the relationship of these four parts and further developments of the subject as well as referencing the papers noted above.

P. P. Ewald

Preface

Several years ago wanting to know more about the transmission of light in materials, I found that an explanation of the electromagnetic theory of transmission of light in materials was contained in the book, "Principles of Optics" by Born and Wolf. There I also found an explanation of the Ewald-Oseen Extinction Theorem and reference to their original papers. In order to have a better understanding of the principles, I attempted to translate Ewald's paper which actually was published in two parts. I found it to be not only basic to the Ewald-Oseen Extinction Theorem but basic to an understanding of the principles of reflection and refraction of light when considered on an atomic level. The basic nature of the paper is the justification for this translation.

In printing this translation, the page format and the placement of headings has followed standards established for AFCRL reports. The numbered equations of the original paper which started with one in each section, have been numbered so that the numbers are consecutive throughout the translation.

The illustrations have been reproduced photographically from those in the original paper, and have also been numbered consecutively. Many of the more complicated mathematical equations have been reproduced photographically to insure against error in transcription. There was considerable use of "italics" in the original paper, but since the reproduction is from typed copy and no italic characters were available, italicized

statements and words are indicated by underlines. Quantities and equations within the text have not been italicized. However, no confusion should arise since the letters in question - namely, a , b , c , l , m , n , etc., were not used with different meanings with normal and mathematical italic type faces in the original paper.

L. M. Hollingsworth

Contents

INTRODUCTION TO PART I AND PART II	1
I. DISPERSION THEORY	7
1. STATEMENT OF PROBLEM	7
2. METHOD OF CALCULATION	8
3. SOME GENERAL REMARKS ON PROCEDURE	15
4. RELATIONSHIP TO OTHER THEORIES	19
5. MATHEMATICAL PREPARATIONS	21
5.1 The Single Integral	21
5.2 The Triple Integral	23
6. THE TOTAL POTENTIAL AND THE OPTICAL FIELD	24
7. SECOND SUMMATION: EXCITATION FIELD	31
8. APPROXIMATIONS DUE TO THE SMALLNESS of α, α_0	37
9. THE DISPERSION FORMULA	40
10. THE EQUATION OF THE NORMAL SURFACE	44
11. THE BIREFRINGENCE D	46

Contents

II. THEORY OF REFLECTION AND REFRACTION	49
1. INTRODUCTION	49
1.1 Problem	49
1.2 Procedure of Computation	50
2. THE SUMMATION OF THE DIPOLE WAVES	52
2.1 Outside Space, z Negative; Pole Series (b)	57
2.2 Inside Space, $z + 2c$ Positive; Pole Series (a) and (c)	58
2.3 The Boundary Region, z Between 0 and $-2c$	59
3. APPROXIMATIONS FOR OPTICAL WAVES	61
3.1 Optical Waves	61
4. FULFILLMENT OF THE DYNAMIC CONDITIONS, THE FRESNEL FORMULAS	66
5. OBSERVATIONS ON THE TRANSITION FROM THE MICROSCOPIC FIELD TO THE VECTORS $\mathbf{E}$ AND $\mathbf{D}$ OF THE MAXWELL THEORY	70
6. WAVE EXT., TOTAL REFLECTION	75
7. TRANSITION TO THE CASE OF X-RAYS	80
ACKNOWLEDGMENTS	84
APPENDIX A. Postscript 1970 by P. P. Ewald	A1
APPENDIX B. Biography of P. P. Ewald	B1

On the Foundations of Crystal Optics

Introduction to Part I (Dispersion Theory) and Part II (Theory of Reflection and Refraction)

The problem of providing a foundation of crystal optics by means of the structure theory of crystals formed the subject of my thesis (Ewald, 1912⁽¹⁾). The task then before me was to investigate whether dipoles, which, since the work of Drude, had been assumed to account for the dispersion of light, would also account for double refraction when suitably arranged in an anisotropic manner so as to serve as a model of a simple orthorhombic crystal. If this were the case, one would furthermore have to determine whether the magnitude of such double refraction is of a comparable order to that observed in crystals of similar values of the axial ratios. These questions arose, to be sure, before the experiments of Laue and his co-workers on the interference of X-rays in crystals had shown the correctness of the structure theory of crystals. It was, in fact, hoped to obtain an argument in favor of this theory from the theory of dispersion.

(Received for publication 26 June 1970)

(1) Reference added 1970: Ewald (1912) Dispersion und Doppelbrechung von Elektronengittern (Kristallen). (Thesis, Univ. of Munich, 1912) Göttingen, Dieterische Universitäts - Buchdruckerei, p 45.

The result of the thesis is that without having to assume extreme ratios of the axes, a lattice arrangement of dipoles can by itself give a strong birefringence because of the inequality of the spacing of the dipoles in the different directions. Furthermore, that the law connecting the index of refraction to the direction of propagation of the wave (illustrated by the surface of wave normals) is fulfilled. The assumed simple model of the crystal is thus correct from the point of view of optics.

The numerically calculated birefringence of a lattice with the same axial ratios as those of anhydrite (anhydrous calcium sulphate), led to different values of the birefringence than the observed ones. This was ascribed to the fact that the crystal consists not of one simple crystal lattice, but of many simple interpenetrating lattices. This type of crystal structure has been revealed by X-ray investigations and is now well established.

It is not yet possible today to decide whether the birefringence in some cases is determined entirely by the anisotropy of positions, or whether in all cases there exists also an anisotropy of the bonding, that is, a different restoring force in the directions of each of the three principal axes. At present we do not know of a birefringent crystal of sufficiently simple chemical nature for which the crystal structure is known from X-ray diffraction so that its positional birefringence can be calculated on the basis of the methods used in my dissertation. Then by comparison with experience, one could evaluate the anisotropy of the bonds and thereby obtain information about the internal forces acting in the crystal. Sulphur might be a suitable example as soon as its structure is entirely known. The numerical calculations which are fully given in the thesis have been considerably shortened in the present publication.

Our main interest lay in the problem of the birefringence: its magnitude and the way it comes about. However it turned out that some novel facets appeared from the point of view of the general theory of dispersion. The precisely given positions of the dipoles made a more rigorous calculation possible than in past treatments of the dispersion of isotropic bodies which required averaging procedures. Small conceptual differences here

forced attention upon themselves. The first question is: How, exactly, are the outgoing fields of the individual dipoles to be summed in the crystal which we think of as being of infinite size? This is a question of convergence which however remains meaningful also for the finite crystal.

A second new aspect, at least with respect to the Planck and Lorentz theory, was the omission of the "incident wave". Planck and Lorentz let the single dipole be excited by an electrical force, which consists of two parts, the "external one" ($\mathcal{E}_0$) and a second contribution which comes from the other dipoles of the body. The first part, $\mathcal{E}_0$, found no place in the theory of the infinite crystal. The mutual interactions alone maintained the dipole vibrations, so that the complete state of oscillation of the crystal had to be looked upon as some type of free-vibration, not requiring an external forcing field. The idealization of the problem by considering a crystal with no boundary was found to contain justification for the omission of the exterior exciting field $\mathcal{E}_0$. In a finite crystal, its boundary should play the role of deleting the incident wave $\mathcal{E}_0$ for the entire interior of the crystal. Only in this way was it possible to understand that in an unbounded crystal no exterior excitation can be noticed.

In this assertion lies the admission that for a complete understanding of dispersion theory the treatment of reflection and refraction is absolutely essential. This supplementary theory of the events in the boundary of the body will be given in the second part of this paper, and that part is new and original.⁽¹⁾ It treats the following problem:

Given a bounded crystal - which fills the entire upper half space above the boundary surface $z = 0$; from below a plane wave $\mathcal{E}_0$ is incident on the boundary at an arbitrary angle. Find the mode of vibration of the dipoles, and whether this leads to the formation of refracted and reflected waves - and implicitly to the Fresnel formulas. Finally, can the anticipated influence of the boundary be shown?

(1) Note added on galley proof. Mr. C. W. Oseen recently, in a different context has also called attention to the necessity to treat dispersion theory in the sense discussed above. Oseen, C. W. (1915) Physikal. Zeitschr. 16:404.

This problem of the bounded crystal appears at first sight to be considerably more difficult than that of the unbounded crystal. In actual fact, however, the vibrational mode as found in the thesis which resembles the free vibrations of a mechanical system is a stepping stone for dealing with the bounded crystal. Just because the surface layer screens the inside of the crystal from the invasion of the external wave, a mode of oscillation inside the crystal becomes possible as if there were neither a boundary nor an incident wave. Since this mode of vibration is well known to us for the unbounded crystal, it requires little effort to reduce the case of the bounded crystal to the case of the unbounded crystal.

The second part of this paper may be considered an extension or continuation of the problem treated in the thesis and in particular as a proof of a statement made there. Furthermore, a particular point that was not treated before will be discussed. This is the transition from the precise microscopic field that exists between the dipoles, to the macroscopic average field that is used in the phenomenological theory for the treatment of reflection and refraction and whose amplitudes are determined by the Fresnel formulas. These are concerned with the comparison of the field strengths outside and inside of the crystal. One has therefore to watch closely the relations between the calculated microscopic field in free space between the dipoles of the body and the Maxwellian field quantities $\mathcal{E}$ and $\mathcal{D}$ inside the body. For the dispersion problem proper, we are only interested in establishing the phase velocities with which the quantities $\mathcal{E}$, $\mathcal{D}$ and $\mathcal{H}$ propagate. For this purpose it is not necessary to distinguish between $\mathcal{E}$ and $\mathcal{D}$, and according to the thesis, the results can be obtained solely from the vector potential without even specifying the field strengths. This discussion, however becomes necessary in connection with the Fresnel formulas. Section 5, Part II, really belongs just as much to Part I (dispersion) as to the theory of refraction, but was omitted from Part I in order not to interrupt the train of thought for the treatment of dispersion theory.

Finally, let it be mentioned that the behavior of crystals for very short waves (x-rays), is contained in the rigorous calculations. See Section 7. It is of particular interest to follow the transition from the case of x-rays with a multiplicity of interference rays to the case of optical rays with only three waves (the incident, the refracted, and the reflected waves).

BLANK PAGE

Part I. Dispersion Theory*

1. STATEMENT OF PROBLEM

(1) Given a simple orthorhombic lattice with lattice points at cartesian coordinates

$$X = 2al; \quad Y = 2bm; \quad Z = 2cn$$

(l, m, n integers from $-\infty$ to $+\infty$)

Electrons are bound to these lattice points with quasi-elastic and isotropic forces. The last statement means that the restoring force between the electron and the lattice point depends only on the magnitude but not on the direction of the displacement. With what velocity and what attenuation can one have a plane wave of a given frequency propagate in this medium?

(2) We begin with the reminder that for the purpose of calculating optical fields the oscillating electrons of charge $-e$ can be replaced by oscillating dipoles. We imagine fixed charges $+e$ and $-e$ to be placed at the lattice points. Then $+e$ together with the displaced electron form the

* Shortened thesis Munich 1912. I wish to not fail to express my sincere thanks in this place to Professor Sommerfeld for his suggestions and support in this work.

dipole, while the remaining charges $-e$ at the lattice points produce an electrostatic field in which we are not interested.

In the equation of a plane wave:

$$\mathcal{E} = \Re(Ee^{-i\omega t + iks}),$$

E is the amplitude, ω the frequency, and s is the coordinate measured in the direction of the wave normal. The sought for values of phase velocity and attenuation are contained in the constant k whose real part k_1 gives the phase velocity $q = \omega/k_1$, while the imaginary part k_2 produces the attenuation.

One now makes the assumption that the planes perpendicular to the planes $s = \text{constant}$ are planes of equal dipole phase. The phase velocity q with which the phase propagates in the lattice is at first not fixed. One now shows that the superposition of the outgoing individual fields of the dipoles produces a (continuous) field strength which has likewise the form of a plane wave; this being a direct consequence of the assumption made for the (discrete) dipole oscillations: furthermore, that the frequency and attenuation of both waves - of the field and of the dipole excitation - are identical. In order to determine these quantities which are both contained in the complex constant k , we have to insure the dynamical possibility or consistency of the original assumption. This requires that each dipole is excited into oscillation by the field impinging on it. This condition can be fulfilled for any frequency by proper choice of the arbitrary constant k and this yields the relationship between frequency and phase velocity, namely the dispersion formula.

2. METHOD OF CALCULATION

Let XYZ be the coordinates of a dipole and xyz the coordinates of a field point; and further let

$$S = X \cos(s, x) + Y \cos(s, y) + Z \cos(s, z)$$

be the distance of a dipole from the plane of equal dipole phase which contains the origin: let q be the (undetermined) dipole phase velocity, n the frequency in the time factor e^{-int} so that

$$e^{-in/q \cdot S}$$

is the phase factor propagating in the direction $-S$ of the dipole excitation.

The individual field emitted by a dipole with the charge $\pm e$ is then characterized by the Hertz vector

$$\mathfrak{H} \cdot e^{-int} = e a \cdot \frac{1}{R} \cdot e^{-in \left(t - \frac{R}{c} + \frac{S}{q} \right)} \quad (1)$$

in which $R = [(X-x)^2 + (Y-y)^2 + (Z-z)^2]^{1/2}$, signifies the distance from dipole to field point. The term R/c in the exponent indicates that the field itself propagates toward the field point with the same velocity c as in free space.*

Since all dipoles have the factor e^{-int} in common, we omit it and introduce

$$\frac{n}{c} = k_0, \quad \frac{n}{q} = k \quad (2)$$

and then write for the significant part of the Hertz vector

$$\mathfrak{H} = e a \cdot \frac{1}{R} \cdot e^{ik_0 R - ik S}, \quad (3)$$

and recall that in a material with dielectric constant 1, the electric field strength $\mathfrak{E}$ and the magnetic field strength $\mathfrak{H}$ can be obtained through the operations

*The original reads "pure ether" but free space has been substituted as the modern equivalent expression.

$$\begin{aligned} \mathcal{E} &= \text{curl curl } \mathfrak{P}; & \mathfrak{H} &= \frac{1}{c} \text{curl } \mathfrak{P} \\ &= \text{grad div } \mathfrak{P} - \Delta \mathfrak{P}. \end{aligned} \quad * \quad (4)$$

According to the equation fulfilled by $\mathfrak{P}$, namely

$$\Delta \mathfrak{P} + k_0^2 \mathfrak{P} = 0 \quad (5)$$

$\Delta \mathfrak{P}$ may be replaced by $-k_0^2 \mathfrak{P}$

The total field is then obtained from the potential

$$\Sigma \mathfrak{P} = e a \Sigma \frac{e^{i k_0 R - i k S}}{R} = e a \Pi, \quad (6)$$

where the sum extends over all the dipoles of the lattice. This total potential becomes infinite as $1/R$ near each dipole and thus in solid materials it is spatially an extremely rapidly changing field since in solid bodies many hundreds of molecules occur within a wavelength of light. For this abruptly changing wave field, the average optical field represents an overall ground swell which can be obtained by averaging.

The average optical field is by no means decisive for the oscillations of the individual dipoles. These, rather, are caused by the excitation field that we write

$$e a \Pi' = \Sigma' \mathfrak{P} \quad (7)$$

where the accent indicates that the sum shall be taken over all the excited dipoles with the exception of the one under consideration. The excitation field is accordingly equal to the total field reduced by the contribution

*Note: The operator Δ is now more often written ∇^2 and this change could have been made in translation, but it was decided to keep the original form.

coming from the test dipole. The equation for the oscillation then is given [from Eq. (4)]

$$m \ddot{p} + g \dot{p} + f p = e \mathcal{E} \\ = e^2 \text{curl curl} (a \Pi' \cdot e^{-i\pi t}). \quad (8)$$

Here

$$p = a \cdot e^{-i\pi t}$$

is the elongation, m is the mass, e is the charge, f is the quasi-elastic force, g the coefficient of friction. By introducing the proper frequency of the dipole

$$\pi_0^2 = \frac{f}{m}$$

the equation becomes

$$a \left(-\pi^2 + \pi_0^2 - i\pi \frac{g}{m} \right) = \frac{e^2}{m} \text{curl curl} a \Pi'.$$

To shorten this we set

$$\frac{m}{e^2} \left(\pi_0^2 - \pi^2 - i\pi \frac{g}{m} \right) = \Omega \quad (9)$$

and replace curl curl with its equivalent: (grad div $-\Delta$). Thus we have to fulfill

$$a (\Omega + \Delta \Pi') = \text{grad div} a \Pi' \quad (10)$$

or in components (the indices on the potential Π' indicate differentiation while those on the amplitude a signify component formation):

$$\begin{aligned}
a_z(\Omega + \Delta\Pi' - \Pi'_{zz}) - a_y\Pi'_{zy} - a_x\Pi'_{zx} &= 0 \\
-a_x\Pi'_{xz} + a_y(\Omega + \Delta\Pi' - \Pi'_{yy}) - a_z\Pi'_{yz} &= 0 \\
-a_z\Pi'_{zz} - a_y\Pi'_{yy} + a_x(\Omega + \Delta\Pi' - \Pi'_{xx}) &= 0.
\end{aligned} \tag{10a}$$

The condition for simultaneous fulfillment of these three homogeneous equations is:

$$\begin{vmatrix}
\Omega + \Delta\Pi' - \Pi'_{xx} & -\Pi'_{xy} & -\Pi'_{xz} \\
-\Pi'_{xy} & \Omega + \Delta\Pi' - \Pi'_{yy} & -\Pi'_{yz} \\
-\Pi'_{xz} & -\Pi'_{yz} & \Omega + \Delta\Pi' - \Pi'_{zz}
\end{vmatrix} = 0. \tag{11}$$

The Π' ... depend on the constant k which determines the dipole phase and this is the same as that which determines the propagation of the optical field. Thus the determinant Eq. (11) expresses the relationship between the index of refraction

$$n = \frac{c}{v} = \frac{k}{k_0} \tag{12}$$

and the quantity Ω , which contains the electron constants and the frequency of the optical wave. (Law of dispersion.)

We obtain the principal indices of refraction of the orthorhombic lattice by assuming the direction of the dipole oscillation to be along one of the three axes; for each of these axes this can be done with two different directions of propagation of the optical wave. We thus obtain in the first instance six principal indices of refraction, which depend on the condition of polarization and on the direction s of the wave normal:

Oscillation	$a = a_x$	Propagation along	$z : k_{xz}, v_{xz}$
	.		$y : k_{xy}, v_{xy}$
	$a = a_y$		$x : k_{yx}, v_{yx}$
	.		$z : k_{yz}, v_{yz}$
	$a = a_z$		$x : k_{zx}, v_{zx}$
	.		$y : k_{zy}, v_{zy}$

Herein lies a departure of the lattice from the actual crystal in which only three principal indices of refraction are observed, and we will now consider the origin of the six indices in a more direct way.

Let the dipole amplitude be in the direction of the x-axis. This leaves us two different ways to dispose of the plane of constant dipole phase (see Figure 1):

(1) The phase plane is the xy-plane (propagation along z).

(2) The phase plane is the xz-plane (propagation along y). On account of the non-equivalence of the y- and z- directions in an orthorhombic lattice, these two cases lead to two different exciting fields and there are in general actually six principal indices of refraction to distinguish.

It is clear, however, without further calculation that the difference between the two indices of refraction ν_{xy} and ν_{xz} is very slight. For it arises from the fact that a phase difference exists between the test dipole and its neighbors - in one case the neighbor in the z-direction, in the other case those along y. The phase difference between adjacent dipoles being very small anyway, because of the large wavelength of the optical field, any effect produced by the inequality of the two phase differences will be extremely small.

It will therefore be sufficient henceforth to distinguish only the three principal refractive indices ν_x , ν_y , ν_z which refer to the directions of oscillation a_x , a_y , a_z respectively. Their differences determine the birefringence. The differences between them are considerably greater than those caused by the directions of wave propagation, because they are nearly

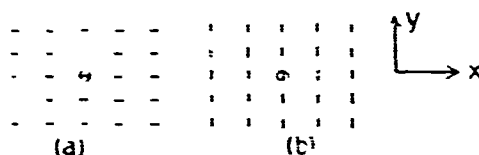


Figure 1

independent of the wavelength. Figure 1a shows the dipoles with their oscillations in the x-direction and Figure 1b the same lattice with the oscillations in the y-direction. For an understanding of the physical significance of the difference between ν_{xy} and ν_{xz} (optical rotation power), see the concluding portion of Section 4.

When computing the principal indices of refraction (propagation along one axis and oscillation along another axis), the three coordinate planes of the lattice remain planes of symmetry and for this reason at the lattice points the differential quotients $\Pi'_{xy} = \Pi'_{xx} = \Pi'_{yz} = 0$; and thus the determinant reduces to the diagonal terms and one obtains from Eq. (10a) the equation for the determination of ν_{xz}

$$\Omega + (\Delta\Pi' - \Pi'_{xx})_z = 0. \quad (13)$$

The index z is to remind one that the specialized formulas for propagation along z must be used in order to find ν_{xz} . In this equation the second derivatives of the potential - as the approximations show - contain only the ratio $k/k_0 = \nu$, but not the frequency n which appears only in the electron constant Ω .

If one eliminates the electron constant Ω , from Eq. (13) and the similar equation for ν_{yz}

$$\Omega + (\Delta\Pi' - \Pi'_{yy})_z = 0 \quad (13a)$$

one obtains an implicit relationship between the two principal indices of refraction ν_{xz} and ν_{yz} , which does not depend on the frequency n ; nor does it depend on the number and characteristic constants of the dipoles; this relationship

$$(\Delta\Pi' - \Pi'_{xx})_z = (\Delta\Pi' - \Pi'_{yy})_z. \quad (14)$$

depends solely on the structure. In the case of an orthorhombic lattice this equation leads to a relationship previously found by T. H. Havelock (see Section 4):

$$\frac{1}{\nu_x^2 - 1} - \frac{1}{\nu_y^2 - 1} = \text{Const.} = D_{xy},$$

whose non-dependence on the wavelength Havelock found to be confirmed in some crystals. The calculation of Π'_{xx} and Π'_{yy} enables us to determine the magnitude of the measure D of the structural birefringence for a given axial ratio.

3. SOME GENERAL REMARKS ON PROCEDURE

(1) The summation of the individual potentials

$$a \frac{e}{R} e^{i(k_0 R - k S)} \quad (15)$$

which is required in the previous section for the calculation of the optical field as well as of the "exciting field," has no definite meaning for real values $k_0 = n/c$ and $k = n/q$. Because of the oscillating factors in the numerator, the decrease of the partial terms as $1/R$ is just about sufficient to make possible a finite value of the sum; however, this value is indefinite and dependent upon the manner of the execution of the summation because the convergence of the series is "conditional". To arrive at an unequivocal result, the observations of the previous sections have to be supplemented by summation instructions which have to be derived from the nature of the physical problems.

To start with, one is relieved of all difficulties of convergence by considering a finite piece of crystal which consists of a large number of dipoles N . When a light wave falls onto this crystal, the dipoles begin to oscillate and create together with the incoming wave $\mathcal{E}_0$, an electric field

in the body. This, in turn, determines the manner of the oscillations, the propagation velocity, etc., as shown in the preceding sections.

In the finite body, this field depends essentially on the boundaries. If the surface on which the incidence occurs is plane, the wave is refracted. If it is undulating, diffraction spectra of different orders develop and all of the diffraction phenomena at the boundaries of the body are within the solution of this problem. The dependence on the surface extends throughout the entire inside of the body.

Obviously, the meaning of the problem of dispersion is more restricted than the proposed task: it is enough to determine the propagation velocity of a wave, that is, a material constant which is independent of outside conditions (boundaries). The very existence of this problem shows that the essential process of wave propagation at a point of the body is caused by the environment of this point. Neither the boundaries nor the outer wave $\mathcal{E}_0$ can influence this process - or else it would have to vary in the body from place to place. The progress of a wave inside of a body is a dynamically closed process which inherently has the capacity that when excited in one part of the body, it can move on to other places without the help of external forces. The variety of the diffraction phenomena of a bounded body results because the boundary initiates many types of oscillations in the body. These then continue to develop in a "dynamically closed" manner.

If our aim is, in the pure theory of dispersion, to find the "dynamically closed" state of oscillation, then we have to eliminate the influence of the boundary. We must, because a material constant is involved, consider only the influence of the environment at the place of the propagation. That the environment should extend everywhere over many wavelengths does not enter into the question of convergence. The simplest way to proceed is to return to the infinite crystal, but to introduce instead of the everywhere equally strong dipole amplitude a a gradually diminishing amplitude at distances far from the test point

$$a e^{-\alpha R}$$

If $\kappa \ll 1/\lambda$, this produces no change within many wavelengths from the test point. However, it assures the determinancy of the problem and the unconditional convergence of the sum of the individual potentials. This sum may now be written

$$\sum a e^{-\kappa R} \frac{e^{ik_0 R - ikS}}{R} = \sum a \frac{e^{ik_0' R - ikS}}{R},$$

and is formally identical with the old sum except for the fact that

$$k_0' = k_0 + i\kappa$$

has the imaginary part κ .

In the case of an absorbing crystal, for which one would have to assume an amplitude growing exponentially in the direction opposite to that of propagation ($k = k_1 + ik_2$) definite values of the sums can evidently only be obtained by assuming a value of κ larger than k_2 . For $\kappa = k_2$ would just compensate for the increase of dipole amplitude and reduce the sums to the same indefiniteness as in the non-absorbing crystal and $\kappa = 0$.

In the following we re-introduce in place of k_0' explicitly

$$k_0 = k_{01} + ik_{02},$$

whereby $k_{02}(=\kappa) > k_2$. The result of the summation with this complex k_0 , is one of the possible sum values with real k_0 . By executing the summation with a complex value and reducing its imaginary part to zero in the sum, a value of the sum is obtained which is free of any influence of the boundary.

(2) From the last statements, it should not be surprising that in the foregoing paragraphs no mention was made of any wave $\mathcal{E}_0$ incident from the outside. The force which stands on the righthand side of the equation for the oscillations contains only fields radiated by the other dipoles of the body, not as with Planck and H.A. Lorentz, a term corresponding to an

incident wave. Such a term would contradict the interpretation of the wave propagation as a "dynamically self-contained" phenomenon. The impossibility of the existence of the primary wave $\mathcal{E}_0$ inside the body is evident from the result of the summation of the individual fields: the assumed dipole oscillations produce a field, which by itself has the form of a plane wave travelling with velocity q . If this field were to be superimposed on the primary field $\mathcal{E}_0$ which has another direction and travels with another velocity c , these fields would never combine within the body to form a single plane wave.

According to the point of view in subparagraph (1), it is clear that the absence of $\mathcal{E}_0$ must be due to an effect of the boundary. The plane boundary of a body fulfills the functions of

- a. compensating the incident wave everywhere in the interior of the body;
- b. initiating the "dynamically self-contained" mode of propagation of the refracted wave;
- c. producing the reflected wave in outer space.

A forceful proof for the required effect of the boundary may be obtained from considering a ray of limited cross section falling under an angle on a plane boundary (Figure 2). But for the fact that the exposed part of the boundary produces a plane wave of equal direction and phase velocity, but opposite phase, as the incident wave, there would have to exist an optical field also in the region between the dashed lines of the figure.

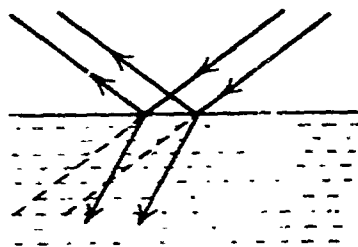


Figure 2

4. RELATIONSHIP TO OTHER THEORIES

This work comes closest to the dispersion theory of Lorentz and Planck. The main difference is that in their treatment of the isotropic body, average values are used whereas in this paper strict consideration of the geometric location of the dipoles is employed.

One further difference of the basic assumptions has already been mentioned above; namely, that we omit the incident wave $\mathcal{E}_0$. It remains to say, however, that the results of the above mentioned theories are confirmed in this paper when the results are specialized for a cubic crystal.

In computing the birefringence, this work exceeds the aims of the above mentioned theories and is closely connected with a group of papers by Lord Rayleigh⁽¹⁾, T. Havelock⁽²⁾, and N. Kasterin⁽³⁾. These papers start out from Rayleigh's investigation of the properties of a medium in which obstacles in the form of small spheres of different dielectric constant are imbedded in a lattice arrangement. Only the static problem of finding the potential distribution is treated. On this basis Havelock uses the dielectric constant for the composite medium and assumes that it varies with wavelength in accordance with Cauchy's formula. This gives him a dispersion formula.

Kasterin, however, has succeeded with a method similar to those used by Rayleigh to find a solution for the wave equation $\Delta u + k^2 u = 0$ which fulfills the boundary condition on the surfaces of the spheres and which represents plane waves. It is a pity that this excellent piece of work has only been applied to acoustics problems.

The connection between Kasterin's work and the present study is based upon the remark⁽⁴⁾, that the way a plane wave is affected by a small

(1) Lord Rayleigh (1892) Phil. Mag. 34.

(2) Havelock, T.H. (1906) Proc. Roy. Soc. London 77:170.

(3) N. Kasterin, Lorentz-Jubelband, or Versl. Acad. Wetensch. Amsterdam 460:1897-1898.

(4) See P. Debye, Der Lichtdruck auf Kugeln, diss. p. 50.

spherical obstacle can be replaced in the first approximation by superposition of a dipole spherical wave issuing from the obstacle.

In a second paper, Havelock⁽¹⁾ has approached birefringence with an idea occurring in Planck's theory, namely that in isotropic bodies there exists around each test dipole a spherical cavity such that the molecules lying within it do not on the average excite the test dipole in the middle of the cavity. According to Havelock, there exists in a crystal an ellipsoid of similar properties. The birefringence can then be expressed by its eccentricity, but it is not connected to measurable quantities. Havelock finds in his work, the relationship D_{av} (Section 2), between the principal refractive indices which should be independent of the wavelength.

Langevin⁽²⁾ comes to the same result by assuming that each molecule is anisotropic. Thus, if it is thought of as being isolated in space, it is assumed to have three different eigen-frequencies. We shall show in Section 9 that the qualitative result of the assumption of an anisotropic molecule must be the same as that resulting from the assumption of a lattice structure.

NOTE: Since the first publication of this paper (thesis), the same problem has also been treated by Max Born in his "Dynamik der Kristallgitter" (Teubner, 1915). Born's mathematical method follows closely the presentation which Hilbert gave of the same subject in his lectures on optics in 1913. In these a consistent mathematical development of potentials according to powers of the small quantities $k, k, \alpha, \beta, \gamma$ etc., is carried out. Born has treated the problem in a very general form, namely for non-orthogonal lattices containing different sorts of electrons or ions. He also assumes that the charges are not bound elastically to fixed lattice points, but only to one another, and that therefore mechanical oscillations occur simultaneously with electrical ones. The case treated in my thesis can be derived from Born's equations by assuming an orthorhombic lattice in which all simple lattices from which it is built have infinite mass except one. The one remaining type of charge oscillates then as though it were bound to fixed lattice points.

(1) Havelock, T. (1907) *Proc. Roy. Soc. London* 80:28.

(2) Langevin, P. (1910) Sur les biréfringences électrique et magnétique, *Le Radium* 7.

The conclusions which Born himself has drawn from his presentation exceed considerably my work in one point; namely in the explanation of the optical rotation. If a beam propagates in the direction of the optical axis, then the polarization direction and the velocity are in the first approximation independent of each other. Only in second order approximation - which would yield in our presentation the difference between v_{xy} and v_{xz} - can one find the anticipated relationship which can, as is well known, be expressed thus: Two circular polarized beams of opposite sense propagate with different velocities.

Born does not furnish numerical calculations which would show what part of the birefringence results from the influence of the lattice arrangement and what part from the bonding.

For the present paper, I have taken from Hilbert's lectures the idea of the consistent development in series of k_0 etc., at that point where agreement of the dispersion formula with the Lorentz-Planck dispersion formula is to be shown in the special case of a cubic lattice. (Section 9).

5. MATHEMATICAL PREPARATIONS

To accomplish the summation of the individual potentials, two representations of

$$\frac{e^{ikR}}{R}$$

prove to be particularly suitable, since they do not contain the irrational value $R = (x^2 + y^2 + z^2)^{1/2}$, but the first or second powers of the coordinates. We speak of one of these as the single, of the other as the triple integral.

5.1 The Single Integral

Let the value k^2 lie in the complex plane within the angle 2ϕ to the real axis.

Then

$$\int e^{-\epsilon^2 + \frac{k^2}{\epsilon^2}} d\epsilon, \quad (16)$$

converges when ϵ is so behaved that for its small values $\Re(\frac{k^2}{\epsilon^2}) < 0$, and for large values $\Re(\epsilon^2) > 0$. This is, for example, the case when ϵ runs out from the origin within the sector $-\pi/4 + \phi$ and $-\pi/4$ and continues in a straight line to infinity.

The value of the integral may easily be determined (see Riemann-Weber, d. part. Dgl. d. math. Physik I, 5th Edition, p. 23):

$$\int e^{-\epsilon^2 + \frac{k^2}{\epsilon^2}} d\epsilon = \frac{\sqrt{\pi}}{2} e^{2ik} \quad (17)$$

By putting $\epsilon = R \cdot \eta$, the path of integration remains essentially unchanged and we have

$$\int e^{-R^2 \eta^2 + \frac{k^2}{R^2} \cdot \frac{1}{\eta^2}} d\eta = \frac{\sqrt{\pi}}{2} \frac{e^{2ik}}{R}$$

Finally by writing $(1/2)k R$ instead of k , we obtain the single integral

$$\frac{e^{ikR}}{R} = \frac{2}{\sqrt{\pi}} \int_{(0)}^{\infty} e^{-R^2 \epsilon^2 + \frac{k^2}{4\epsilon^2}} d\epsilon, \quad (18)$$

whose path can be deformed so as to coincide with the positive real axis, except near the origin, where it has to lie within the angle ϕ . (See Figure 3).

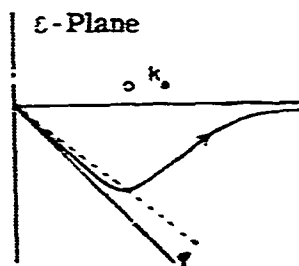


Figure 3

5.2 The Triple Integral

In order to obtain a representation in which the exponent of e is a linear function of x, y, z , the simplest way is to transform the single integral representation Eq. (19) as follows:

We have

$$\int_{-\infty}^{+\infty} e^{-\frac{k^2}{4t^2} - i k x} d\lambda = e^{-x^2 t^2} \int_{-\infty}^{+\infty} e^{-\left(\frac{1}{2t} + i x t\right)^2} d\lambda = 2t e^{-x^2 t^2} \cdot \sqrt{\pi}. \quad (19)$$

We substitute in Eq. (18) in place of $e^{-x^2 t^2}$ the integral over λ standing on the left hand side of this equation and similarly for $e^{-y^2 t^2}$ and $e^{-z^2 t^2}$ analogous integrals over the variables μ and ν . Thus we obtain the quadruple integral:

$$\frac{e^{i k x}}{R} = \frac{2}{\sqrt{\pi}} \cdot \left(\frac{1}{2\sqrt{\pi}} \right)^3 \int \int \int \int e^{-\frac{k^2 + \mu^2 + \nu^2}{4t^2} - i(\lambda x + \mu y + \nu z) + \frac{k^2}{4t^2}} d\lambda d\mu d\nu e^{-3} dt \quad (20)$$

The order of integrations may be changed and the integration over t performed first:

$$\int_0^\infty e^{-\frac{k^2 + \mu^2 + \nu^2}{4t^2}} e^{-3} dt = -\frac{2}{k^2 - \mu^2 - \nu^2 - \nu^2}. \quad (21)$$

This leaves the triple integral standing

$$\frac{e^{i k x}}{R} = -\frac{1}{2\pi^3} \int \int \int_{-\infty}^{+\infty} e^{i k x + i \mu y + i \nu z} \frac{d\lambda d\mu d\nu}{k^2 - \mu^2 - \nu^2 - \nu^2}; \quad (22)$$

which integral has the desired very suitable form for summation.

It is seen that the substitution $-\lambda$, .. in place of $+\lambda$, .. changes nothing but the sign of the exponent. Furthermore, if R is the distance of the point xyz from the point XYZ , rather than from the origin, the generalized integral is valid

$$\frac{e^{ikR}}{R} = -\frac{1}{2\pi^3} \iiint e^{\pm i\lambda(X-x) \pm i\mu(Y-y) \pm i\nu(Z-z)} \frac{d\lambda d\mu d\nu}{k^2 - \lambda^2 - \mu^2 - \nu^2} \quad (23)$$

Finally it should be noted that it is possible to shift the integration path from the real axes provided that the condition

$$k^2 - \lambda^2 - \mu^2 - \nu^2 = 0$$

is avoided. In particular, if the positive imaginary part of k is nonzero, identical integration paths a short distance above the real axis may be employed for the three variables. (That is, between the axis and the point k).

6. THE TOTAL POTENTIAL AND THE OPTICAL FIELD

In order to obtain the total potential of the orthorhombic point lattice, we first make use of the triple integral [Section 5, Eq. (23)] and sum the individual potentials

$$\frac{e^{i(k_0 E - kS)}}{R} = -\frac{e^{-ikS}}{2\pi^3} \iiint \frac{e^{\pm i\lambda(X-x) \pm i\mu(Y-y) \pm i\nu(Z-z)}}{k^2 - \lambda^2 - \mu^2 - \nu^2} d\lambda d\mu d\nu \quad (24)$$

The coordinates of the lattice points are:

$$X = 2al, \quad Y = 2bm, \quad Z = 2cn,$$

where l , m , n take on all integral positive and negative values. (a , b , c) are the half distances between neighboring points on the crystal axes and,

at the same time, the spacings of adjacent planes of symmetry in the crystal. k_0 is the propagation constant of the individual spherical wave, k that of the dipole phase. We write

$$\begin{aligned} kS &= k(X \cos(S, x) + Y \cos(S, y) + Z \cos(S, z)) \\ &= \alpha X + \beta Y + \gamma Z; \end{aligned} \quad (25)$$

whereby

$$\alpha^2 + \beta^2 + \gamma^2 = k^2. \quad (25a)$$

The total potential then is

$$\begin{aligned} \Pi &= -\frac{1}{2\pi^2} \sum_{l, m, n}^{\dots, +\infty} \iiint e^{2ial(\pm l - \alpha) + 2ibm(\pm m - \beta) + 2icn(\pm n - \gamma)} \\ &\quad \cdot e^{\mp i l x \mp i m y \mp i n z} \frac{d\lambda d\mu d\nu}{k_0^2 - \lambda^2 - \mu^2 - \nu^2}. \end{aligned} \quad (26)$$

We carry out the summation under the integral. The triple sum is then the product of three linear geometric series of which the first is

$$\sum_{l=-\infty}^{+\infty} e^{2ial(\pm l - \alpha) \mp i l x}.$$

Convergence of this sum for positive and negative values of l can be achieved thanks to the uncertainty of the sign of λ , provided that the imaginary part k_{02} is greater than k_2 (see Section 3). For if this is so, we can let λ run on a line parallel to the real axis and above α , similarly for μ and ν with imaginary parts greater than those of β and γ^* respectively. Thus $+\lambda - \alpha$ has a positive imaginary part which insures that the sum converges for positive values of l . The sum over negative l converges when we reverse the sign of λ .

*One finds the proof in the thesis note 1.

The summation over l therefore yields

$$\begin{aligned}
 \sum_{l=-\infty}^{+\infty} \dots &= \sum_0^{-\infty} e^{2ial(\lambda-a)-i\lambda x} + \sum_{-1}^{-\infty} e^{-2ial(\lambda+a)+i\lambda x} \\
 &= \frac{e^{-i\lambda x}}{1 - e^{2ia(\lambda-a)}} + \frac{e^{+i\lambda x + 2ia(\lambda+a)}}{1 - e^{2ia(\lambda+a)}} \\
 &= \frac{e^{-i\lambda x}(1 - e^{2ia(\lambda+a)}) + e^{i\lambda x + 2ia(\lambda+a)}(1 - e^{2ia(\lambda-a)})}{1 + e^{4ia\lambda} - 2\cos 2a\alpha e^{2ia\lambda}} \quad (27) \\
 &= \frac{e^{-i\lambda(x+2a)} - e^{-i\lambda x + 2ia\alpha} + e^{i\lambda x + 2ia\alpha} - e^{i\lambda(x+2a)}}{2\cos 2a\lambda - 2\cos 2a\alpha} \\
 &= -i \frac{\sin \lambda(x+2a) - e^{2ia\alpha} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} .
 \end{aligned}$$

Appropriately performing the summations over m and n , the Total Potential thus is

$$\begin{aligned}
 \Pi = -\frac{i}{2\pi^2} \iiint & \frac{\sin \lambda(x+2a) - e^{2ia\alpha} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} \cdot \frac{\sin \mu(y+2b) - e^{2ib\beta} \sin \mu y}{\cos 2b\mu - \cos 2b\beta} \\
 & \cdot \frac{\sin \nu(z+2c) - e^{2ic\gamma} \sin \nu z}{\cos 2c\nu - \cos 2c\gamma} \cdot \frac{d\lambda d\mu d\nu}{k_0^2 - \lambda^2 - \mu^2 - \nu^2}, \quad (28)
 \end{aligned}$$

where the integration paths for $(\lambda \mu \nu)$ are lines parallel to the real axes, and lying above $(\alpha \beta \gamma)$, respectively. (Path 1, Figure 4).

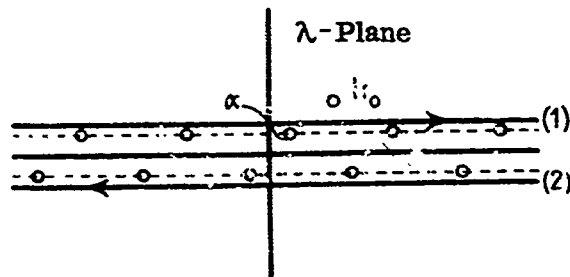


Figure 4

Let us remove a part of the integrand. We write in the λ -integral

$$\begin{aligned} & \frac{\sin \lambda(x+2a) - e^{2ia} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} \\ &= \frac{\sin \lambda x \cos 2a\lambda + \cos \lambda x \sin 2a\lambda - e^{2ia} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} \\ &= \sin \lambda x + \frac{\cos \lambda x \sin 2a\lambda - e^{2ia} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} \end{aligned}$$

Here the first term, $\sin \lambda x$, has lost the denominator and with it a pole of the integrand. For this part of the integral we can then shift the path of λ to the real axis. Because $\sin \lambda x$ is an odd function and $k_0^2 - \lambda^2 - \mu^2 - \nu^2$ is an even function of λ , the integral is zero. The same is true for the proposed integrations over μ and ν . The integral then is:

$$\begin{aligned} II = -\frac{i}{2\pi^2} \iiint & \frac{\cos \lambda x \sin 2a\lambda - e^{2ia} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} \dots \\ & \dots \frac{d\lambda d\mu d\nu}{k_0^2 - \lambda^2 - \mu^2 - \nu^2} \end{aligned} \quad (29)$$

It is easy to convert this integral into a sum built up from the residues at the poles of the integration. The poles for the λ -integration are given by

$$\cos 2a\lambda_0 = \cos 2a\alpha, \quad \lambda_0 = \frac{l\pi \pm a\alpha}{a},$$

where l is a positive or negative integer. In the λ -plane the poles lie on the two parallels to the real axis which go through the point $\pm a$. (Figure 4).

The integrand is an odd function of λ . Then by the substitution $\lambda' = -\lambda$, the horizontal original path (1) is replaced by path (2) which lies symmetrically to (1) below the lower row of poles and is traversed from $+\infty$ to $-\infty$. Adding a factor $1/2$ for integration in each variable, we may use the path (1) + (2), thus obtaining

$$\Pi = -\frac{i}{16\pi^2} \underbrace{\int \int \int}_{(1)+(2)} \frac{\cos \lambda x \sin 2a\lambda - i \sin \lambda x \sin 2a\alpha}{\cos 2a\lambda - \cos 2a\alpha} \dots$$

$$\dots \frac{d\lambda d\mu d\nu}{k_0^2 - \lambda^2 - \mu^2 - \nu^2};$$

The path (1) + (2) is a path closed at infinity which can be contracted to circulation of the row of poles, the circulation occurring in the negative sense.

The residue at one of the poles $\lambda_0 = (l\pi \pm a\alpha)/a$ is obtained by development of the integrand denominator in terms of $(\lambda - \lambda_0)$:

$$\frac{\cos \lambda_0 x \sin 2a\lambda_0 - i \sin \lambda_0 x \sin 2a\alpha}{-2a \sin 2a\lambda_0 (\lambda - \lambda_0) + \dots}$$

$$= \frac{1}{\lambda - \lambda_0} \cdot \frac{-1}{2a} \left(\cos \lambda_0 x - i \sin 2a\alpha \cdot \frac{\sin \lambda_0 x}{\sin 2\lambda_0 a} \right) + \dots$$

$$= \frac{1}{\lambda - \lambda_0} \cdot \frac{-1}{2a} \left(\cos \lambda_0 x - i \sin \lambda_0 x \cdot \frac{\sin 2a\alpha}{\pm \sin 2a\alpha} \right) + \dots$$

$$= \frac{1}{\lambda - \lambda_0} \cdot \frac{-1}{2a} e^{\mp i \frac{l\pi \pm a\alpha}{a} x} + \dots,$$

where the upper signs refer to the poles of the upper set and conversely.

The sum of all residues which result from the integration over λ is

$$\frac{2\pi i}{2a} \left\{ \sum_{l=-\infty}^{+\infty} l \frac{e^{-i \frac{l\pi + a\alpha}{a} x}}{k_0^2 - \mu^2 - \nu^2 - \left(\frac{l\pi + a\alpha}{a} \right)^2} + \sum_{l=-\infty}^{+\infty} l \frac{e^{+i \frac{l\pi - a\alpha}{a} x}}{k_0^2 - \mu^2 - \nu^2 - \left(\frac{l\pi - a\alpha}{a} \right)^2} \right\}$$

or since the second sum is identical with the first, (differing only by having $-l$ in place of l for the summation index):

$$\frac{2\pi i}{a} \sum_{l=-\infty}^{+\infty} l \frac{e^{-i \frac{l\pi + a\alpha}{a} x}}{k_0^2 - \mu^2 - \nu^2 - \left(\frac{l\pi + a\alpha}{a} \right)^2}.$$

By the same procedure regarding the other integrations we obtain the total potential:

$$\Pi = -\frac{\pi}{2abc} \sum_{l, m, n} \frac{e^{-i\frac{l\pi + a\alpha}{a}x - i\frac{m\pi + b\beta}{b}y - i\frac{n\pi + c\gamma}{c}z}}{k_0^2 - \left(\frac{l\pi + a\alpha}{a}\right)^2 - \left(\frac{m\pi + b\beta}{b}\right)^2 - \left(\frac{n\pi + c\gamma}{c}\right)^2}. \quad (30)$$

This series which is equivalent to the results Eqs. (28) and (29) is a precise description of the electromagnetic vector potential which results from the superposition of the spherical waves.

By placing the factor

$$e^{-i(ax + by + cz)} = e^{-iks}$$

before the summation sign, the potential has the form

$$\Pi = -\frac{\pi}{2abc} \cdot e^{-iks} \cdot \Sigma, \quad (30a)$$

where the remaining sum Σ is a periodic function (Fourier series) with the development range $\pm a, \pm b, \pm c$. The factor e^{-iks} differs from the factor for the dipole phase only therein, that in place of the discontinuous variable S (the coordinate of a dipole) the continuous coordinate s is substituted.

In order to obtain a clear picture of the potential given by Eq. (30a), one must keep in mind, that the development range of the periodic function is extremely small in comparison with the distance that the factor e^{-iks} requires for a full period. The distance between the atoms in the body are of the order of 10^{-8} cm while a light wave, even in strongly refractive material, is not smaller than 10^{-5} cm. There are thus 100 - 1000 atoms to a light wave length.

The factor e^{-iks} thus hardly changes within an elementary parallelepiped. In order to visualize Eq. (30a) we have primarily to think of a periodic function of very complicated character repeating around each atom. After that, proceeding from atom to atom, we have to modify the amplitude of the function according to the modulation given by the slowly changing factor.

For the optical wave, roughness and spikes of the microscopic field are immaterial and the periodic function, in its repetition from atom to atom, contributes only by its average value. This is given by the constant term 000 of the Fourier series, namely

$$\frac{1}{k_0^2 - k^2}$$

Therefore the potential of the optical field is:

$$a \bar{\Pi} = a \cdot \frac{\pi}{8abc} e^{-iks} \frac{1}{k^2 - k_0^2}; \quad (31)$$

completed by the time factor e^{-int} it describes a plane wave in the direction -s whose exponent contains k as the wave constant and not k_0 as do the individual spherical waves. Accordingly, the optical potential satisfies the wave equation

$$\Delta \bar{\Pi} + k^2 \bar{\Pi} = 0.$$

The wave constant k, which was originally assumed, and the ratio

$$v = \frac{k}{k_0}$$

have to be determined from the "dynamic equilibrium condition" (Eq. 10, Section 2).

The factor $(k_0^2 - k^2)^{-1}$ determines the amplitude a of the oscillating electrons for a given intensity of the observed optical wave. For small values of $(k_0 - k)$, namely when the index of refraction is not far from one, the amplitudes are particularly small. This is fundamentally the reverse of the statement, that slightly resonating electrons have only a small influence on the index of refraction.

7. SECOND SUMMATION: EXCITATION FIELD

The form of the total potential as derived in the previous paragraphs, shows clearly the origin of the optical field. In particular, it was easy to see that the total potential has the same phase e^{iks} at all lattice points as that of the dipoles. This means that in the calculation of dipole oscillation (Eq. 10, Section 2) we may restrict ourselves to the dipole at the origin. For it we have to find the excitation potential, or rather its second derivative, which by means of the determinant (Eq. 11, Section 2), determines the phase constant k .

It is not possible to separate the effect of the test dipole 000 from the sum Eq. (30) of the previous paragraph. This is due to the fact that the sum becomes infinite for $x = y = z = 0$ by non-uniform convergence. The contribution which has to be subtracted, $R^{-1} e^{ik_0 R}$, diverges in an entirely different manner. It is not possible to combine both terms before they become infinite. Furthermore, the form in which Π appears does not allow double differentiation.

It is very fortunate, especially with regard to the numerical calculations, that the summation also may be performed with the help of the single integral for the individual potential [Eq. (18), Section 5]. One obtains a single integral over a product of theta functions and exponential functions and one can effectively perform the subtraction of the divergent terms from the integrand so that in this calculation of the excitation potential, no indefinite term $\infty - \infty$ occurs. This is fundamentally the merit of the

ϑ -transformation theorem, by virtue of which each ϑ -function admits of two entirely different developments. In one development, each term can be interpreted by us as the effect of a dipole, and therefore the contribution of the test dipole to the total potential can be recognized and subtracted. In the other development, the contributions of the individual dipoles are completely mixed up and each term of the theta-expansion matches one of the terms from which in the previous sections we have built up the total potential in the form of the triple sum. By suitable use of both descriptions, it is not difficult to subtract the contribution of the test dipole, while retaining the essential parts of the fields of the other dipoles.

We proceed from the integral Eq. (18), Section 5:

$$\frac{e^{ikR}}{R} = \frac{2}{\sqrt{\pi}} \int_{(0)}^{\infty} e^{-R^2 s'^2 + \frac{k^2}{4s'^2}} ds'; \quad (32)$$

in which we substitute a dimensionless variable s by setting

$$s = \frac{2d}{\sqrt{\pi}} s', \quad x_0^2 = \frac{e^2 k^2}{\pi}, \quad x^2 = \frac{e^2 k^2}{\pi}, \quad P^2 = \frac{\pi}{4} \cdot \frac{x^2 + y^2 + z^2}{e^2}, \quad (33)$$

where d is a length of the order of magnitude of a , b , c which we will choose later as the geometric mean of the lengths of these sides. Then

$$\frac{e^{ikR}}{R} = \frac{1}{d} \int_{(0)}^{\infty} e^{-\frac{\pi}{4} \frac{P^2}{e^2} \{ (X-x)^2 + (Y-y)^2 + (Z-z)^2 \} + \frac{x_0^2}{e^2}} d\epsilon. \quad (34)$$

The condition of Section 5, regarding the direction under which the path of integration has to leave the origin is not affected by this transformation.

We multiply by the dipole phase factor and sum: This gives for the total potential

$$\Pi = \frac{1}{d} \sum_{l, m, n} \int_0^\infty e^{-\frac{\pi}{4} \frac{r^2}{d^2} \{ (2al - r^2 + (2bm - y^2 + (2cn - z^2)) - 2iaml - 2ib\beta m - 2ic\gamma n + \frac{r^2}{d^2} \}} d\epsilon. \quad (35)$$

The part of the exponent which is dependent on the summation index l is:

$$- \epsilon^2 \frac{a^2}{d^2} l^2 + \left(\epsilon^2 \frac{a^2}{d^2} - 2ia\alpha \right) l.$$

The simplest theta function (which is referred to also as ϑ_{∞} or ϑ_3) admits of the following series expansion:

$$\vartheta(\tau | i\tau) = \sum_{n=-\infty}^{+\infty} e^{-n^2\tau - 2\pi i n \tau}. \quad (36)$$

By interchanging the sequence of summation and integration, we have therefore in Eq. (35) before us

$$\vartheta \left(\frac{a^2}{\pi} + i \frac{a^2}{2d^2} \epsilon^2 \middle| i \frac{a^2}{d^2} \epsilon^2 \right)$$

and the total potential is

$$\begin{aligned} \Pi = \frac{1}{d} \int_0^\infty & \vartheta \left(\frac{a^2}{\pi} + i \frac{a^2}{2d^2} \epsilon^2 \middle| i \frac{a^2}{d^2} \epsilon^2 \right) \cdot \vartheta \left(\frac{b^2}{\pi} + i \frac{b^2}{2d^2} \epsilon^2 \middle| i \frac{b^2}{d^2} \epsilon^2 \right) \\ & \cdot \vartheta \left(\frac{c^2}{\pi} + i \frac{c^2}{2d^2} \epsilon^2 \middle| i \frac{c^2}{d^2} \epsilon^2 \right) \cdot e^{-\frac{\pi}{4} \frac{r^2}{d^2} + \frac{r^2}{d^2}} d\epsilon. \end{aligned} \quad (37)$$

If we desire to drop the dipole 000 from the summation, we omit the 000 term and we obtain the excitation potential:

$$\Pi = \frac{1}{d} \int_0^{\infty} \left\{ \vartheta \cdot \vartheta \cdot \vartheta - 1 \right\} e^{-\frac{t^2}{\varepsilon} + \frac{\pi^2}{\varepsilon^2}} dt. \quad (38)$$

In order to perform the integration, we utilize the series expansion of ϑ . The original series Eq. (36) converges quickly when ε is large; we obtain a useful expansion for small ε through the ϑ -transformation formula⁽¹⁾.

$$\begin{aligned} \vartheta(v; i\tau) &= \frac{1}{\sqrt{i}} e^{-\frac{\pi^2}{\tau}} \vartheta\left(i\frac{\tau}{i} \middle| \frac{i}{\tau}\right) \\ &= \frac{1}{\sqrt{i}} e^{-\frac{\pi^2}{\tau}} \sum e^{-\pi^2 \frac{\tau}{i} - 2\pi i \frac{\tau}{i}}. \end{aligned} \quad (39)$$

By means of the 'transformed' and the 'original' series, the following integrals for Π are produced:

$$\Pi = \begin{cases} \frac{1}{d} \cdot \frac{d^3}{abc} \int \sum e^{-\frac{\pi^2}{\varepsilon} \left\{ \left(i\frac{d}{a} + \frac{da}{\varepsilon} \right)^2 + \dots \right\} - i\pi \left\{ \left(i\frac{d}{a} + \frac{da}{\varepsilon} \right) \frac{\varepsilon}{d} + \dots \right\} + \frac{\pi^2}{\varepsilon}} d\varepsilon \\ \frac{1}{d} \int \sum e^{-\pi^2 \left\{ \left(i\frac{\varepsilon}{d} + \frac{\varepsilon}{2d} \right)^2 + \dots \right\} - 2i\pi \left\{ \left(i\frac{\varepsilon}{d} + \frac{\varepsilon}{2d} \right) + \dots \right\} + \frac{\pi^2}{\varepsilon}} d\varepsilon. \end{cases} \quad (40)$$

The integrands are identically equal for each value of ε . We will abbreviate the exponents in the first and second integral as I and II respectively. Then the derivatives are:

(1) See for example, A. Krazer, *Lehrbuch der Thetafunktionen*, p. 96, Teubner.

$$\frac{\partial^2 \Pi}{\partial x^2} = \Pi_{xx} \quad \text{and} \quad \frac{\partial^2 \Pi}{\partial x \partial y} = \Pi_{xy}:$$

$$\Pi_{xx} = \begin{vmatrix} -\frac{\pi^2}{abc} \int \sum \left(l \frac{d}{a} + \frac{d\epsilon}{\pi} \right)^2 \cdot e^l \epsilon^{-3} d\epsilon \\ \frac{1}{d^2} \int \sum \left\{ \pi^2 \epsilon^4 \left(l \frac{a}{d} + \frac{x}{2d} \right)^2 - \frac{\pi^2 \epsilon^2}{2} \right\} e^u d\epsilon. \end{vmatrix} \quad (41)$$

$$\Pi_{xy} = \begin{vmatrix} -\frac{\pi^2}{abc} \int \sum \left(l \frac{d}{a} + \frac{d\epsilon}{\pi} \right) \left(m \frac{d}{b} + \frac{d\beta}{\pi} \right) e^l \cdot \epsilon^{-3} d\epsilon \\ \frac{1}{d^2} \int \sum \pi^2 \epsilon^4 \left(l \frac{a}{d} + \frac{x}{2d} \right) \left(m \frac{b}{d} + \frac{y}{2d} \right) e^u d\epsilon. \end{vmatrix} \quad (42)$$

We must now subtract from Π_{xx} the second derivative with respect to x of the individual potential of the test dipole at the origin. For this purpose, we divide the integration for Π_{xx} in two parts: from 0 to E and from E to ∞ . The crossover place E remains at present indeterminate. On the first part of the path we use the series with the exponent I, on the second, the series with II. The quantity to be subtracted is [see Eq. (34), this Section]:

$$\frac{\partial^2}{\partial x^2} \frac{e^{ik_0 L}}{R} = \frac{1}{d^2} \int_0^\infty \left(\pi^2 \epsilon^4 \frac{x^2}{4d^2} - \pi \frac{\epsilon^2}{2} \right) e^{-\pi^2 \epsilon^2 + \frac{\pi^2}{d^2}} d\epsilon. \quad (43)$$

Again, we divide this integral in two parts at the crossover point E , and find that the integral from E to ∞ is identical with the term 000 of the second half of Π_{xx} . The subtraction can therefore be achieved by putting an accent on the sum sign in order to indicate that the term 000 has to be omitted. If we simultaneously take the value at $x = y = z = P = 0$ we obtain

* As published, due to typographical error, the last power within the brackets was given as 8.

$$\begin{aligned}
\Pi'_{xx} &= \Pi_{xx}^{(1)} + \Pi_{xx}^{(2)}; \\
\Pi_{xx}^{(1)} &= -\frac{\pi^2}{abc} \sum_{\alpha} \int_0^E \left(l \frac{d}{a} + \frac{d\alpha}{\pi} \right)^2 \cdot e^l \cdot \varepsilon^{-3} d\varepsilon + \frac{1}{d^2} \int_0^E \frac{\pi \varepsilon^2}{2} e^{\frac{x_0^2}{\varepsilon^2}} d\varepsilon \\
\Pi_{xx}^{(2)} &= \frac{1}{d^2} \sum_{\alpha} \int_0^E \left\{ \left(l \pi \varepsilon^2 \frac{a}{d} \right)^2 - \frac{\pi \varepsilon^2}{2} \right\} e^l d\varepsilon,
\end{aligned} \tag{44}$$

where the exponents at the origin are:

$$\begin{aligned}
I &= -\frac{\pi}{\varepsilon^2} \left\{ \left(l \frac{d}{a} + \frac{d\alpha}{\pi} \right)^2 + \dots \right\} + \frac{x_0^2}{\varepsilon^2}, \\
II &= -\pi \varepsilon^2 \left\{ l^2 \frac{a^2}{d^2} + \dots \right\} - 2i(l\alpha a + \dots) + \frac{x_0^2}{\varepsilon^2}.
\end{aligned}$$

In $\Pi_{xx}^{(1)}$ the general sum term may easily be integrated when one takes into consideration that $\varepsilon^{-3} d\varepsilon = \frac{1}{2} d(1/\varepsilon^2)$. The indefinite integral produces no contribution at $\varepsilon = 0$ because of the direction of the departure from the zero point; there remains the value at the upper limit E and one has:

$$\begin{aligned}
\Pi_{xx}^{(1)} &= \frac{\pi^2}{2abc} \sum_{\alpha} \frac{\left(l \frac{d}{a} + \frac{d\alpha}{\pi} \right)^2 e^{-\frac{\pi}{E^2} \left(l \frac{d}{a} + \frac{d\alpha}{\pi} \right)^2 + \dots} + \frac{x_0^2}{E^2}}{x_0^2 - \pi \left(\frac{d}{a} + \frac{d\alpha}{\pi} \right)^2 - \dots} \\
&\quad + \frac{\pi}{2d^2} \int_0^E e^{\frac{x_0^2}{\varepsilon^2}} \varepsilon^2 d\varepsilon.
\end{aligned} \tag{45}$$

Here one sees plainly the relation to the total potential in the form of Section 6. For $E = \infty$, Π'_{xx} goes over into the (formally) differentiated total potential Eq. (30), Section 6, taken at the origin and diminished by the (likewise infinite) test dipole term. If E is, however, given a finite value, then all the terms are finite because the infinite effect of the test dipole has been taken out of $\Pi_{xx}^{(2)}$.

Through a similar calculation the following is obtained:

$$\begin{aligned}\Pi_{xy} &= \Pi_{xy}^{(1)} + \Pi_{xy}^{(2)} \\ \Pi_{xy}^{(1)} &= \frac{\pi^2}{2abc} \sum \frac{\left(l \frac{d}{a} + \frac{d\alpha}{\pi}\right) \left(m \frac{d}{b} + \frac{d\beta}{\pi}\right) \cdot e^{-\frac{\pi}{k^2} \left\{ \left(l \frac{d}{a} + \frac{d\alpha}{\pi}\right)^2 + \dots \right\} + \frac{\pi_0^2}{k^2}}}{\pi_0^2 - \pi \left(l \frac{d}{a} + \frac{d\alpha}{\pi}\right)^2 + \dots} \\ \Pi_{xy}^{(2)} &= \frac{\pi^2}{d^3} \frac{ab}{d^2} \int_E^\infty \sum' l \cdot m \cdot e^{\frac{d}{a} l} \cdot e^{\frac{d}{b} m} d\epsilon.\end{aligned}\quad (46)$$

In contrast to $\Pi_{xx}^{(1)}$, there exists no term in $\Pi_{xy}^{(1)}$ which corresponds to the single potential to be subtracted. The reason for this is that in the differentiation of Eq. (34) with respect to x and y no term appears which becomes infinite independently of the direction. In spite of the factor $l \cdot m$ the accent on the sum in $\Pi_{xy}^{(2)}$ may not be omitted, because the integral over the term $l = m = n = 0$ diverges.

The series Eq. (45) for $\Pi_{xx}^{(1)}$ and (44) for $\Pi_{xx}^{(2)}$ give the necessary data for the calculation of the principal indices of refraction inasmuch as it is possible to find them without the use of approximations. In the following sections, we shall make use of the simplifications resulting from the actual orders of magnitude in the case of propagation of visible light in solid bodies.

8. APPROXIMATIONS DUE TO THE SMALLNESS OF π , π_0 .

Equations (42) and (43) established in Section 7 for the derivatives of the excitation potential at the place of the dipole 000 can be considerably simplified when π_0 , π and $d \cdot \alpha$, $d \cdot \beta$, $d \cdot \gamma$ are small numbers. This is always the case when visible light in solid material is treated, due to the fact that d is of the same order as the interatomic distances and k_0 , k , α , β , γ are of the order of $1/\lambda$. As already mentioned, the ratio

of d/λ is of the order of $1/100$ to $1/1000$ and the above quantities are of the same order.

Under these circumstances, terms such as $d\alpha/\pi$ occurring in the series Eqs. (44) and (45) Section 7, play a role only when $i = m = n = 0$.

We omit, therefore, in all terms of the series, the small values after having segregated the term 000. This term is in $\Pi_{xx}^{(1)}$ according to Eq. (14), Section 7:

$$\frac{\pi^2}{2abc} \frac{d^2 \alpha^2}{\pi^2} \frac{1}{\kappa_0^2 - \kappa^2} = \frac{\pi}{2abc} \frac{\nu^2 \cos^2(S, x)}{1 - \nu^2},$$

where $\nu = v/\kappa_0$ is the index of refraction. The term is therefore of the order of magnitude 1 in comparison to κ^2 .

Because of the smallness of ν , the integral appearing in Eq. (45) of the preceding section

$$\int_{(0)}^E \epsilon^2 d\epsilon$$

has an integrand which differs appreciably from ϵ^2 only where the integrand is very small anyhow. We commit an error at most of the order κ_0^2 , when we replace this integral with one independent of κ_0 :

$$\int_{(0)}^E \epsilon^2 d\epsilon = \frac{E^3}{3}.$$

After these steps, the first part of the excitation potential becomes

$$H_{xx}^{(1)} = \frac{\pi}{2abc} \left(\frac{\nu^2}{1 - \nu^2} \cos^2(S, x) + \psi_{xx}^{(1)} \right),$$

where in this approximation $\psi_{xx}^{(1)}$ depends neither on x and x_0 nor on the direction of propagation of the wave, but only on a , b , and c .

Similar considerations are applied to $\Pi_{xx}^{(2)}$ Eq. (44), Section 1. Also here the small quantities of the order of x_0^2 can be disregarded in the exponent of the e-function wherever they stand next to terms of the order of magnitude 1. Only for the term 000 would this not be the case; but this term is excluded because of the accent. Thus $\Pi_{xx}^{(2)}$ becomes entirely independent of the small quantities. We write in this approximation

$$\Pi_{xx}^{(2)} = \frac{\pi}{2abc} \psi_{xx}^{(2)}$$

and obtain altogether for the excitation potential, after setting $\psi^{(1)} + \psi^{(2)} = \psi$:

$$\Pi'_{xx} = \frac{\pi}{2abc} \left(\frac{r^2}{1-r^2} \cos^2(S, x) + \psi_{xx} \right) \quad (47)$$

and similarly

$$\Pi'_{xy} = \frac{\pi}{2abc} \left(\frac{r^2}{1-r^2} \cos(S, x) \cos(S, y) \right). \quad (48)$$

In Π'_{xy} and analogous derivatives, it is evident that no term ψ_{xy} appears; for this term would be independent of the direction of propagation and thus would also exist in the case of propagation along one of the coordinate axes. But we know from symmetry that in this case $\Pi'_{xy} = 0$ (see above).

Finally, we have in this notation:

$$\Delta \Pi' = \frac{\pi}{2ab}, \left(\frac{r^2}{1-r^2} + \Delta \psi \right). \quad (49)$$

9. THE DISPERSION FORMULA

In Section 2 we had to distinguish six cases of principal indices of refraction, depending on the directions of the oscillation and of the ray. But we have previously suspected that pairs of these indices would so nearly coincide, that their discrimination would be practically superfluous. As a matter of fact, we can now write down the equation for the propagation of the optical wave in the z direction with oscillation in the x direction:

$$\Omega + (\Delta \Pi' - \Pi'_{xx})_{\alpha=\beta=0, \gamma=k} = \Omega + \frac{\pi}{2abc} \left(\frac{\nu_x^2}{1-\nu_x^2} + \Delta\psi - \psi_{xx} \right) = 0$$

and we find, that in this approximation, because of the independence of ψ from the direction of the ray, the identical equation would be valid for ν_{xy} , that is for $\alpha = \gamma = 0$, $\beta = k$. We therefore omit the distinction between ν_{xy} and ν_{xz} , and we write ν_x for the principal index of refraction for dipole oscillations parallel to the x axis.

Let v be the volume $8abc$ which contains one dipole; then for the index of refraction we have the equation

$$\frac{\nu_x^2}{\nu_x^2 - 1} = \frac{v}{4\pi} \Omega + \Delta\psi - \psi_{xx} \quad (50)$$

or

$$\nu_x^2 = 1 + \frac{1}{\frac{v}{4\pi} \Omega + \Delta\psi - \psi_{xx} - 1}. \quad (51)$$

We recall the significance of the electron constant Ω [Eq. (9), Section 2]

$$\Omega = \frac{m}{e^2} \left(n_0^2 - n^2 - i n \frac{g}{m} \right);$$

For a transparent crystal we set the frictional coefficient $g = 0$. Then the dispersion formula is:

$$v_x^2 = 1 + \frac{1}{\frac{v}{4\pi} \frac{m}{e^2} n_0^2 - (1 - A\psi + \psi_{xx}) - \frac{v}{4\pi} \frac{m}{e^2} n^2} \quad (52)$$

$$= 1 + \frac{1}{B_x - A n^2}; \quad (53)$$

or, introducing the wavelength $\lambda = \frac{2\pi c}{n}$:

$$v_x^2 = 2 + \frac{\lambda^2}{\lambda^2 B_x - 4\pi^2 c^2 A}. \quad (54)$$

We have here before us a dispersion formula with two constants, of the same type as those of H. A. Lorentz* and of Planck. The latter would be in our terminology

$$v^2 = 1 + \frac{1}{\frac{v}{4\pi} \frac{m}{e^2} n_0^2 - \frac{1}{3} - \frac{v}{4\pi} \frac{m}{e^2} n^2},$$

whereas Lorentz has $-(1/3 + s)$ in place of $-1/3$ where s remains indeterminate. In all three formulas this term indicates the influence of the dipoles of the body on a test dipole. The term $1/3$ arises from the contributions of the more distant surroundings of the test dipole, where the dipoles may be spread out into a homogeneous medium. In Lorentz' terms is added the strong but unpredictable influence that near neighbor dipoles exert on the test dipole. For a cubic arrangement of neighboring molecules, H. A. Lorentz proves that $s = 0$, so that in this case his formula becomes identical with Planck's.

*Cited under Section 4.

Also from our equations specialized for a cubic lattice the term $1/3$ is obtained. For in that case one has, by symmetry

$$1 - \Delta\psi + \psi_{xx} = 1 - \frac{1}{3} \Delta\psi. \quad (55)$$

But $\Delta\psi$ can be determined in the following manner: one develops the total potential Eq. (30), Section 6 in powers of k obtaining

$$\Pi = -\frac{4\pi}{v} \left(\frac{1}{k_0^2 - k^2} + [k^{-1} \dots + \dots] \right)$$

when only the lowest powers are written out.

In order to obtain the excitation potential,

$$\frac{e^{ikR}}{R} = \frac{1}{R} + ik - \dots$$

should be subtracted from this. The term with k^{-1} in the development of Π goes over unchanged in the development of Π' . Thus one has

$$\begin{aligned} \Delta\Pi &= -k_0^2\Pi = \frac{4\pi}{v} \left(\frac{k_0^2}{k_0^2 - k^2} + [k \dots + \dots] \right) \\ &= \frac{4\pi}{v} \left(\frac{1}{1 - v^2} + [k \dots + \dots] \right) \\ &= \frac{4\pi}{v} \left(\frac{v^2}{1 - v^2} + 1 + [k \dots] \right). \end{aligned}$$

By comparison with Eq. (49) Section 8 it follows that

$$\Delta\psi = 1,$$

and thus from Eq. (55)

$$1 - \Delta\psi + \psi_{xx} = \frac{1}{3}.$$

Insofar as the isotropic body may be replaced by a cubic arrangement, our formula thus entirely agrees with the Lorentz-Planck formula.

Discussion of the dispersion formula thus obtained would add nothing new regarding the dependence on the wavelength of the indices of refraction and absorption than what has already been discussed by Planck.

Only those details which are produced through the anisotropy of Eq. (52) shall be noted here:

(1) From two measurements of ν_x the quantities A and B_x [Eq. (53)] can be computed. The quantities ν are functions only of the lattice constants a, b, c . Thus by using the value of a single principal index of refraction at two different wave lengths one can obtain the complete dispersion formula not only for this index, but also for the two other principal indices of refraction.

(2) The constant A , independent of which index of refraction would be used in its determination should have the value $\nu/4\pi \cdot m/e^2$ which is characteristic for the lattice. The same value of A should therefore be found, no matter from which principal index of refraction it is obtained. But one finds this relationship not fulfilled. For example for anhydrite, (anhydrous calcium sulphate)

$$A_a = 1.6125 \cdot 10^{-33}; \quad A_\beta = 1.6726 \cdot 10^{-33}; \quad A_\gamma = 1.5501 \cdot 10^{-33},$$

values which deviate by 8 percent.

One of the reasons for this deviation is the fact that the electrons having the proper frequency ν_0 do not form the simple Bravais lattice that has been presupposed here. A second reason is that in addition to the electrons with the characteristic frequency ν_0 others with different eigen frequencies and possibly ions with eigen frequencies in the infrared are contained in the crystal.

The two-constant Eq. (52) is then to be understood as an approximation to a formula of complicated structure and the constant A of the dispersion formula consequently loses its simple meaning and can vary with the

direction of the propagation of the wave or of the polarization. In this connection the observation of Drude⁽¹⁾ is relevant: if in calcite incidence is at right angles to the optical axis, the ordinary ray involved apparently five electrons, and the extraordinary ray two electrons per molecule.

(3) From Eq. (52) it follows immediately that the effect of the lattice arrangement can be interpreted also as an anisotropic bonding of the electrons in the molecules. The variation in the formula of dispersion in passing from ν_x to ν_y , or from ϵ_{xx} to ϵ_{yy} , respectively can be replaced, as Eq. (52) shows, for all wavelengths by a suitable modification of the eigen frequency ν_0 . Located within the lattice, a dipole, which in itself is bound isotropically, acquires through the influence of the structure three different eigen frequencies. The two views, namely attributing the origin of birefringence either to the inherent anisotropy of the ultimate particles, or to the anisotropic arrangement of isotropic particles, thus cannot differ in their qualitative consequences.

10. THE EQUATION OF THE NORMAL SURFACE

We have dealt up to now exclusively with the principal indices of refraction, six of which we had to distinguish originally. With neglect of the terms of the order $(\alpha/\lambda)^2$ we recognized that indices referring to the same polarization are equal to each other. This made the properties of the lattice equal to those of a bi-axial crystal.

The question is whether this equivalence is true for all directions, that is, whether the relationship between the direction of the wave normal S and the index of refraction, which is contained in the determinant (Eq. 11, Section 2), is the same as in the orthorhombic crystal.

Retaining the same degree of accuracy, we introduce the values of Eqs. (47), (48), and (49) from Section 8 in the first line of the determinant.

(1) Ann. d. Phys. 14:691, 1904.

This line becomes

$$\Omega + \frac{\pi}{2abc} \left(\frac{r^2}{1-r^2} (1 - \cos^2(s, x)) + \Delta\psi - \psi_{xx} \right);$$

$$- \frac{\pi}{2abc} \frac{r^2}{1-r^2} \cos(s, x) \cos(s, y); - \frac{\pi}{2abc} \frac{r^2}{1-r^2} \cos(s, x) \cos(s, z).$$

We introduce the abbreviations:

$$N = \frac{r^2}{1-r^2}, \quad \cos(s, x) = \cos 1 \dots$$

and

$$\frac{2abc}{\pi} \Omega + \Delta\psi - \psi_{xx} = \frac{r^2}{1-r^2} = N_x$$

Thus the determinant reads:

$$\begin{vmatrix} N \sin^2 1 - N_x & -N \cos 1 \cos 2 & -N \cos 1 \cos 3 \\ -N \cos 1 \cos 2 & N \sin^2 2 - N_y & -N \cos 2 \cos 3 \\ -N \cos 1 \cos 3 & -N \cos 2 \cos 3 & N \sin^2 3 - N_z \end{vmatrix} = 0.$$

We write in the diagonal $(1 - \cos^2 \dots)$ in place of $\sin^2 \dots$ and take out the factor which is non-zero except in trivial cases:

$$-N^3 \cos^2 1 \cdot \cos^2 2 \cdot \cos^2 3$$

There remains

$$\begin{vmatrix} \left(1 - \frac{N-N_x}{N \cos^2 1}\right) & 1 & 1 \\ 1 & \left(1 - \frac{N-N_y}{N \cos^2 2}\right) & 1 \\ 1 & 1 & \left(1 - \frac{N-N_z}{N \cos^2 3}\right) \end{vmatrix} = 0$$

or, except for a factor:

$$\frac{N \cos^2 1}{N - N_1} + \frac{N \cos^2 2}{N - N_2} + \frac{N \cos^2 3}{N - N_3} - 1 = 0.$$

This equation is nothing other than that of the normal surface of a bi-axial crystal. Replacing the 1 by $\cos^2 1 + \cos^2 2 + \cos^2 3$ and the $N \dots$ by their expression in ν , one obtains the equation in its usual form:

$$\frac{\nu_x^2 \cos^2 1}{\nu^2 - \nu_x^2} + \frac{\nu_y^2 \cos^2 2}{\nu^2 - \nu_y^2} + \frac{\nu_z^2 \cos^2 3}{\nu^2 - \nu_z^2} = 0.$$

Thus we see, that a point lattice of the kind considered, behaves with respect to all directions as a crystal and that from the standpoint of optics there is no objection to even the simplest lattice model of a crystal.

11. THE BIREFRINGENCE D

Already the general considerations of Section 2 have indicated a relationship between the principal indices of refraction, which according to Eq. (14) Section 2, is

$$(\Delta \Pi' - \Pi'_{zz})_z = (\Delta \Pi' - \Pi'_{yy})_z. \quad (56)$$

The indices z on the parentheses indicate that on both sides of the equation the potentials are those for the direction of propagation along z , but that on the lefthand side, polarization is along x and the refractive index is ν_x , while on the righthand side polarization is along y and the refractive index is ν_y .

If we insert in this relationship the approximate values of Eqs. (47), (48), and (49) from Section 8, there results

$$\frac{\pi}{2abc} \left(\frac{v_x^2}{1-v_x^2} + \Delta\psi - \psi_{xx} \right) = \frac{\pi}{2abc} \left(\frac{v_y^2}{1-v_y^2} + \Delta\psi - \psi_{yy} \right)$$

or since Δv is the same in both cases, we obtain

$$\frac{v_x^2 - v_y^2}{(1-v_x^2)(1-v_y^2)} = \psi_{xx} - \psi_{yy} = D_{xy}. \quad (57)$$

The dimensionless quantity on the left, which was recognized already by T. H. Havelock as independent of the wavelength, we find dependent only on the ratios a:b:c of the lattice periods. We call it the measure of structural birefringence.

Through the series Section 7, we are in position to compute these constants for arbitrary ratios a:b:c. For a suitably chosen crossover point E, the series converge rapidly enough to make the numerical calculation easy.

But at present it is not yet promising to undertake the calculation for a known crystal in the hope of finding agreement between calculated and observed values. None of the crystals whose structure is known up to now (Fall 1915) has a structure so simple that one could assume with some measure of certainty that a single kind of electron is adequate to render its optical properties.

The numerical computation and the extension of the theory to more complicated space lattices shall therefore be omitted here.

Only in order to give a general idea about the expected order of the magnitude of the effects, we take for the calculation the ratios a:b:c = 0.8932 : 1 : 1.008 which are the same as those of anhydrite. Calculation leads to

$$D_{xy} = 0.17848 \quad D_{xx} = 0.17983 \quad D_{yy} = 0.00133,$$

whereas the observed values for the same material are:

$$D_{\rho_1} = 0.05145 \quad D_{\rho_2} = 0.05925 \quad D_{\rho_3} = 0.00854$$

From this it may be concluded that the anisotropic arrangement alone is sufficient for explaining the order of magnitude of the observed birefringence.

That the simple assumptions concerning the space lattice do not suffice for rendering the phenomena, is also indicated by the fact that the observed values of D are found independent of the wavelength only to a moderate extent. Thus in quartz the variation of D between the lines C and E of the spectrum is 15 times smaller than the variation of the double refraction $(\omega - \epsilon)$ in the customary sense. In most crystals the variation of D is some 3 to 6 times smaller than that of the usual birefringence, but there are crystals for which D varies more than the birefringence.

Whether the constancy of D could be valid for arbitrarily complex point systems constructed of one kind of electrons may remain undecided. Certainly it is not valid for complex point systems with different sorts of ions, since it is not possible to eliminate the two or more electron constants Ω_i from the two dispersion equations for ν_x and ν_y .

Unfortunately no similar simple relationship for two sorts of electrons or ions could be found between the three principal indices of refraction, which would be valid regardless of wavelength.

Part II. Theory of Reflection and Refraction*

1. INTRODUCTION

1.1 Problem

Let the upper half space be filled with a lattice of dipoles with the coordinates

$$X = 2al; \quad Y = 2bm; \quad Z = 2cn, \quad (58)$$

where l, m, n are integers and l, m assume all positive and negative values, while n assumes only positive values (including 0). Further let a plane wave be incident on the lattice from below and let its wave normal have the direction s_0 (see Figure 1). Its periodic function is

$$e^{-i\pi\left(t - \frac{z}{c}\right)} \quad (59)$$

Find the mode of oscillation of the lattice and the thereby created optical field.

* Continuation of the work by the same title in Ann. d. Phys. 49:1-38, 1916, quoted as Part I.

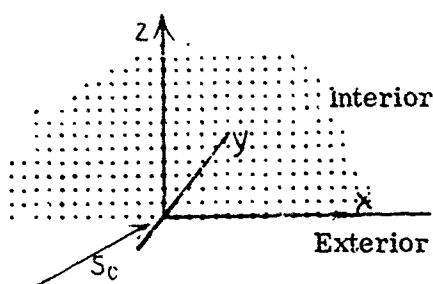


Figure 5

1.2 Procedure of Computation

A procedure which appears to hold little promise would be to write down methodically the excitation equations for all dipoles and to solve them beginning with those at the boundary. The force terms appearing on the right hand side of these equations would themselves depend on all of the dipole oscillations and at that, in a most complicated manner. Consequently the computation of the force without previous knowledge of the dipole oscillations is by no means practical.

Another method of solution exists in making a suitable assumption for the dipole oscillations. We can then actually sum the electric forces and satisfy ourselves that the dynamic conditions are fulfilled for the assumed mode of oscillations.

Therefore, the assumption is made here that in the crystal, up to the very boundary, there exist planes of equal dipole phase. In the unbounded crystal such an assumption leads to the propagation of a plane optical wave. Let the wave normal of this dipole wave be s , and its velocity q .

Then the summation of the spherical waves issuing from the dipoles does not yield, as in the case of the unbounded crystal, only the field of a plane wave, which propagates in the same manner (velocity q , direction s), as the dipole phase. Rather, there occurs besides this field inside the crystal, a plane wave with the velocity c whose direction of propagation s_0 is related to s by the law of refraction. Moreover, in outside space, a

wave comes into existence which is a mirror image with regard to the surface of the last-mentioned internal wave and it has likewise the velocity c . The outside wave is the reflected wave. The first inside wave (q) is the refracted wave. The wave on the inside of velocity c is the one which is required by dispersion theory in order to cancel the incident wave. It begins with full strength at the boundary and can therefore compensate the incident wave s_0 throughout the interior of the crystal. (It is called for short, the internal boundary wave.)

This result already offers the solution of the dynamical conditions for the dipole oscillations. For when this same boundary wave and the incident wave cancel each other inside the crystal, then only the same field remains inside the bounded crystal as in the unbounded crystal, and we only have to see to it that the assumed dipole oscillations are compatible with this field. This is accomplished when the propagation velocity of the refracted wave has the correct value, as was shown in the dispersion theory.

Amplitude and phase of the internal boundary wave are determined by the requirement that the incident wave be cancelled. Since the three waves which are generated by the dipole oscillations show definite amplitude and phase relationships, the amplitudes and phases of the reflected and refracted waves are also determined by this condition of cancellation. The resulting relationship between the amplitudes of the three waves must be that given by the Fresnel formulas provided that the correct transition is made from the microscopic field to the phenomenological Maxwell field.

Let it be said at once that we restrict our thinking to the case of a cubic crystal which is thus also optically isotropic. In this way complications will be avoided, which would only detract from the essentials of the theory. Notwithstanding this, we distinguish in the formulas, the three axis intercepts a, b, c , which, if desired, immediately may be considered as equal.

2. THE SUMMATION OF THE DIPOLE WAVES

As in the first part, Dispersion Theory, we employ the Hertz potential of a dipole in the form of expression Eq. (23), Part I, Section 5, namely with omission of the factor $e a \cdot e^{-i \omega t}$:

$$\frac{e^{i k_0 R}}{R} = -\frac{1}{2\pi^2} \iiint e^{\pm i\lambda(X-z) \pm i\mu(Y-y) \pm i\nu(Z-s)} \frac{d\lambda d\mu d\nu}{k_c^2 - \lambda^2 - \mu^2 - \nu^2} \quad (50)$$

(X, Y, Z) are the coordinates of the dipole, (x, y, z) those of the field point. The variables λ, μ, ν run parallel to the real axis between it and the point k_0 , (which, initially, is thought of as having a positive imaginary part) (Figure 6).

The potentials to be summed are:

$$a \frac{e^{-i(\omega t - k_0 R - iS)}}{R} \quad (61)$$

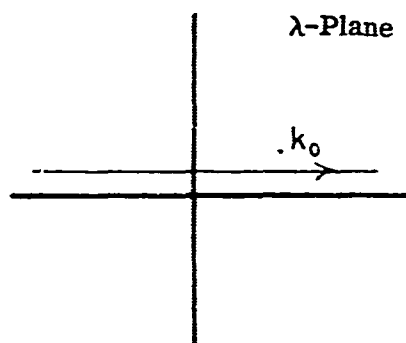


Figure 6

Here (in the terminology of Part I, Section 10)

$$\begin{aligned} kS &= k \cos 1 \cdot X + k \cos 2 \cdot Y + k \cos 3 \cdot Z \\ &= \alpha X + \beta Y + \gamma Z. \end{aligned} \quad (62)$$

This corresponds to a progressive excitation along $+s$.

On performing the summation under the integral sign in Eq. (60), the threefold sum is decomposed into the product of three simple sums. Of these, the first two whose indices l and m run from $-\infty$ to $+\infty$, are identical with the sums over l and m which appear in the corresponding problem of the unbounded crystal (see Part I, Section 6). Our present assumption differs only insignificantly from that made there, namely in the signs of α, β, γ . These are easily changed in the results of the summations. The sum over l thus produces here in the integrand the factor

$$-i \frac{\sin \lambda(x+2a) - e^{-2ia\alpha} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} \quad (63)$$

and similarly for the sum over m .

On the other hand, the summation over the index n extends only from 0 to ∞ and we have:

$$\sum_{n=0}^{\infty} e^{2icn\nu - i\pi z + 2icy\cdot n} = \frac{e^{-i\pi z}}{1 - e^{2ic(\nu - \gamma)}} \quad (64)$$

In order that the sum converge, $\nu + \gamma$ must have a positive imaginary part; that is, ν must be conducted on a path above $-\gamma$. In the former sums, as also in the unbounded crystal, it was required that λ be led above α , μ above β . The feasibility of this requirement provided the sums are unconditionally convergent has been discussed in the thesis and has been attributed to the fact that k_0 has a larger imaginary part than has k .

This assumption is also here necessary for the summation. In the result we go without further discussion to real k , k_0 (see Part I, Section 3).

The total potential of the "half crystal" consequently becomes:

$$\begin{aligned}
 &= \frac{1}{2\pi^3} \iiint \frac{\sin \lambda(x+2a) e^{-2ia\alpha} \sin \lambda x}{\cos 2a\lambda - \cos 2a\alpha} \\
 &\quad \cdot \frac{\sin \mu(y+2b) e^{-2ib\beta} \sin \mu y}{\cos 2b\mu - \cos 2b\beta} \\
 &\quad \cdot \frac{e^{-ivz}}{1 - e^{2ic(v+z)}} \cdot \frac{d\lambda d\mu dv}{k_0^3 - \lambda^2 - \mu^2 - v^2}.
 \end{aligned} \tag{65}$$

From this description, which is valid for all values of (x, y, z) we proceed to a description of Π as a sum by reducing the three integrations to circulations of the poles of the integrand in the complex plane and evaluating these by residues. This entails of itself a distinction of inner and outer space with each containing different groups of waves.

The first two integrations over λ and μ are identical with those in the expression Eq. (23), Part I, Section 6 for the potential of the unbounded crystal. They will be treated just as there. By utilizing the fact that the integrand is an odd function of the variable, the path of Figure 5 can be completed by one running below the real axis. The series of poles of the integrand now lies between the two paths at the positions

$$\lambda_0 = \frac{l\pi}{a} \pm \alpha. \tag{66}$$

Residue development in the integral over l yields the sum

$$\frac{2\pi i}{a} \sum_{-\infty}^{+\infty} \frac{e^{-i(l\pi - a\alpha)x/a}}{k_0^3 - \left(\frac{l\pi - a\alpha}{a}\right)^2 - \mu^2 - v^2} \tag{67}$$

and the same treatment is applied to the integration over μ . (Compare Part I, Section 6.)

By this transformation of the integrations over λ and μ the potential of the bounded crystal becomes:

$$\Pi = - \frac{1}{2ab} \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \int e^{i(l\pi - a\alpha)x/a + i(m\pi - b\beta)y/b} \cdot \frac{e^{-i\gamma z}}{1 - e^{2i\gamma(v+\gamma)}} \frac{d\gamma}{k_0^2 - \left(\frac{l\pi - a\alpha}{a}\right)^2 - \left(\frac{m\pi - b\beta}{b}\right)^2 - \gamma^2} \quad (68)$$

This is a mixed form of the potential, inasmuch as two integral signs have been replaced by summation signs, while the third integration still remains. This form shares with Eq. (65) the advantage that it is valid for the entire space.

We now proceed to replace the last integral also by a sum of residues.

To this end we observe that the integrand has two classes of poles according to whether the first or second denominator is zero.

$$a) 1 - e^{2i\gamma(v+\gamma)} = 0; \quad \gamma = \frac{n\pi}{c} - \gamma, \quad \text{where } n \text{ is an integer} \quad (69)$$

This series of poles corresponds completely to the poles Eq. (68) in the λ and μ plane, at which the residues were taken. The poles lie on a straight line parallel to the real axis and below i , but above the point $-k_0$. This line thus is between the integration path for γ and the point $-k_0$ (Figure 7).*

$$b) \quad \underline{k_0^2 - \left(\frac{l\pi - a\alpha}{a}\right)^2 - \left(\frac{m\pi - b\beta}{b}\right)^2 = \gamma^2},$$

that is

** (70)

$$\gamma_{lm} = \pm \sqrt{k_0^2 - \left(\frac{l\pi - a\alpha}{a}\right)^2 - \left(\frac{m\pi - b\beta}{b}\right)^2}.$$

* The position of this line was stated incorrectly when published and is corrected here.

** The words "that is" have been supplied in translation in order to clearly separate the two statements of Eq. (70).

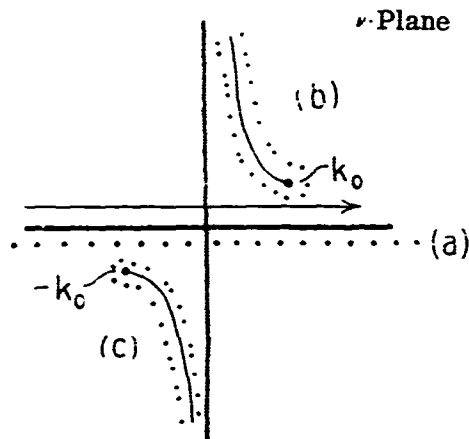


Figure 7

By $\sqrt{\nu}$ we understand the value of the root which is situated in the first quadrant. These poles form a chain which sheathes the branch cut of the root extending from k_0 to infinity*. With increasing l and m the poles become ever more imaginary.

(c) Also the other root value, thus under retention of the positive sign for the principal value, the value

$$\underline{\nu = -\nu_{im}}$$

causes the denominator to vanish.

Altogether: three lines of poles exist in the ν plane as shown in Figure 7.

The integration path for ν runs, as mentioned, parallel to the real axis between the pole series (a) and the series (b). This path can either be deformed upwards and then contracted about the pole series (b) with circulation about each pole in the positive direction: or it can be pulled downward, namely in case that the integrand vanishes exponentially at

* Clearly this branch cut does not exist for the integration over ν since in the denominator only ν^2 appears.

infinity in the lower half plane. In doing so it remains held by the pole series (a) and (c), (both with negative circulation of the poles).

The condition for the upward deformation is that the factor

$$\frac{e^{-i\pi z}}{1 - e^{2i\pi(z+\gamma)}}$$

vanish at infinity in the upper half plane, thus

$$z < 0$$

This is fulfilled in all outer space.

The downward deformation on the other hand is permissible because of the same factor, provided the denominator becomes infinite more rapidly than does the numerator. That is when

$$z + 2c > 0$$

which is true in the entire inside space.

These limits show that a domain exists at the boundary

$$-2c < z < 0$$

in which both deformations are permitted. The boundary domain continues the crystal beyond the cut off at $z = 0$ into outer space and establishes thus the transition from inner to outer space.

We have now to obtain the residues.

2.1 Outside Space : Negative Pole Series (b)

We have

$$\frac{1}{k^2 - \left(\frac{ix - ax}{a}\right)^2 - \left(\frac{mx - bx}{b}\right)^2 - r^2} = -\frac{1}{r - r_+} \cdot \frac{1}{r - r_-}$$

and thus the sum of the residues:

$$\Pi_a = \frac{\pi i}{2ab} \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} e^{-i(l\pi - a\alpha)z - i(m\pi - b\beta)y/b - i\nu_{lm}z} \cdot \frac{1}{1 - e^{2i(c(\nu_{lm} + \gamma))}} \cdot \frac{1}{\nu_{lm}}. \quad (71)$$

This is the potential Π_a which is valid in the outer space. It is a double sum, wherein each term appears to have the form of an ordinary plane wave. In fact, this is true only for the term 00 given that for visible light k_0 , α, β, γ are small compared with $1/a$, $1/b$, $1/c$. For the same reason all other terms for which ν_{lm} is positive imaginary represent inhomogeneous waves. (See, however, Section 7).

2.2 Inside Space. $z + 2c$ Positive: Pole Series (a) And (c)

The residue at a pole of the series (a) is found by development of the integrand in proximity to the pole. There

$$r = r_a = \frac{n\pi - c\gamma}{c}$$

and nearby

$$\frac{e^{-i\nu z}}{1 - e^{2i(c\nu - \gamma)}} = \frac{e^{-i\nu_a z}}{-2ic(r - r_a) + \dots}$$

Because of the negative sense of circulation, the residue is multiplied by the factor $-2\pi i$. The potential becomes — we designate it Π_q :

$$\Pi_q = -\frac{\pi}{2abc} \sum_l \sum_m \sum_n \frac{e^{-i(l\pi - a\alpha)z - i(m\pi - b\beta)y/b - i(n\pi - c\gamma)z}}{i_l^2 - \left(\frac{l\pi - a\alpha}{a}\right)^2 - \left(\frac{m\pi - b\beta}{b}\right)^2 - \left(\frac{n\pi - c\gamma}{c}\right)^2}. \quad (72)$$

Except for the already frequently mentioned change of the signs of α, β, γ this is identical with the potential of the unbounded crystal. [See Part I, Section 6, Eq. (30)].

Finally, as a second part of the inner potential, the double sum over the pole series (c) is to be added; but for the sign of ν_{lm} and a further (-1) because of the negative circulation this is equal to the sum of the residues at the poles (b):

$$\Pi_c = \frac{i\pi}{2ab} \sum_l \sum_m e^{-i(l\pi - \alpha)z} e^{-i(m\pi - \beta)y} e^{i\nu_{lm}} = \frac{1}{1 - e^{-2i\pi(\nu_{lm} - \gamma)}} \cdot \frac{1}{\nu_{lm}}. \quad (72)$$

What has been said about the character of the sum Π_a is valid also for Π_c .

The total inner potential is obtained by adding together the two constituents:

$$\Pi_i = \Pi_a + \Pi_c. \quad (74)$$

2.3 The Boundary Region, z Between 0 And $-2c$

In order to become familiar with these sums we may study the boundary region in which both expressions Π_a and Π_c should be valid. Without going back to the integrals and the deformation of their paths, it is easily shown on the series themselves that both expressions are identical inside the boundary zone.

To this end we assert

$$\Pi_a - \Pi_c = \Pi_b, \quad (74a)$$

and have to prove that the threefold sum on the right-hand side is equal to the difference of the double sums.

It is easily seen that the third summation, over n on the right-hand side has its origin in the Fourier development of the parts containing ν_{lm} in the double sums of the left-hand side. In comparing Eqs. (71), (72) and (73) one sees that for z between 0 and $-2c$ one should have identically

$$\frac{1}{\nu_{lm}} \left(\frac{e^{-i\nu_{lm}z}}{1 - e^{2ic(\nu_{lm} + \gamma)}} - \frac{e^{+i\nu_{lm}z}}{1 - e^{-2ic(\nu_{lm} - \gamma)}} \right) = \frac{i}{c} \sum_{(n)} \frac{e^{-in\pi z/c + i\gamma z}}{\nu_{lm}^2 - (n\pi/c - \gamma)^2}. \quad (75)$$

The easiest way to prove this relationship is to prepare the sum on the righthand side in the following way: Start with the development of $e^{-i\nu z}$ in the interval $0 \dots -2c$ which is

$$e^{-i\nu z} = \frac{1}{c} \sum_{(n)} e^{-in\pi z/c} \frac{\sin \nu c \cdot e^{i\nu c}}{\nu - n\pi/c}. \quad (76)$$

Bring the factor $\sin \nu c \cdot e^{i\nu c}$ over to the other side; there remains a development with coefficients

$$A_n = \frac{1}{\nu_{lm} - n\pi/c}.$$

The coefficients of the series in Eq. (75) are

$$B_n = \frac{1}{\nu_{lm}^2 - (n\pi/c - \gamma)^2} = \frac{1}{2\nu_{lm}} \left\{ \frac{1}{\nu + \gamma - n\pi/c} - \frac{1}{-\nu + \gamma - n\pi/c} \right\}.$$

Therefore the value of the series in Eq. (75) is the difference of two exponential functions:

$$\sum_{(n)} = \frac{e^{i\gamma z}}{2\nu} \cdot c \left\{ \frac{e^{-i(\nu + \gamma)(z+c)}}{\sin(\nu + \gamma)c} - \frac{e^{-i(-\nu + \gamma)(z+c)}}{\sin(-\nu + \gamma)c} \right\}.$$

The righthand side of this equation, completed by a factor i/c is after some easy manipulation equal to the lefthand side of Eq. (75). This shows the identity of both developments, namely Eq. (71) and Eq. (72) + Eq. (73) within the boundary region.

3. APPROXIMATIONS FOR OPTICAL WAVES

Before we proceed to satisfy the dynamic conditions at each dipole by means of the incident optical wave, we will attempt to obtain a survey of the nature of the fields represented by the potentials. There are two cases to distinguish:

(1) In the case of optical waves, the quantity a/λ , that is, the half distance of the lattice points measured in wavelengths, is a small number of the order of magnitude $1/100$. In our equations this means that $k = n/q = 2\pi/\lambda$ and together with it k_0 , α, β, γ are all small compared to $1/a$.

(2) If an X-ray goes through the lattice, the rigorous computation based on the assumption Eq. (61) is valid as in case Eq. (58). This time, however, ak is of the order of magnitude of 10, since the wavelength of the usual X-rays (maximum $1\text{\AA}$) is more than 1000 times smaller than the wavelength of light. This case will be treated in Section 7.

3.1 Optical Waves

3.1.1 THE INNER WAVE Π_q .

The potential Π_q is already known from Part I. It is sufficient here to recall that it contains two factors, one of which has the form e^{iks} of a plane wave which proceeds in the same manner as the dipole phase of the lattice. The other factor is a threefold sum which describes in the manner of Fourier a function which repeats from one elementary parallelepiped to another. The optical field must obviously be defined by an averaging process in such a way that the variations of this rapidly changing constituent are

eliminated. For this purpose, one integrates the second factor over a parallelepiped thus obtaining an average potential

$$\overline{\Pi}_q = \Pi_q^{(0)} = -\frac{\pi}{2abc} e^{-i(ax + \beta y + \gamma z)} \cdot \frac{1}{k_0^2 - k^2}. \quad (77)$$

To see how this is related to the macroscopic field strength, see Section 5.

3.1.2 THE INNER WAVE Π_c

The nature of the terms of this double sum is determined by the value of

$$\nu_{lm} = + \sqrt{k_0^2 - \left(\frac{l\pi - a\alpha}{a}\right)^2 - \left(\frac{m\pi - b\beta}{b}\right)^2} \quad (78)$$

In the optical case only the term (0, 0)

$$\nu_{00} = \sqrt{k_0^2 - (\alpha^2 + \beta^2)} \quad (79)$$

can be real: All other terms are definitely imaginary. All terms of the sum except for $l = m = 0$ represent inhomogeneous waves, whose amplitudes decrease in the z direction according to the factor

$$e^{i\nu_{lm}z}$$

The higher the indices, so much faster is the decrease. If we have

$$\nu_{01} = \sqrt{k_0^2 - \alpha^2 - (\pi/b - \beta)^2} \sim i\pi/b \quad (79a)$$

the factor becomes

$$e^{-\pi z/b}.$$

If we take the case of cubic lattice, where $a = b = c$ it thus occurs that the amplitude of the wave through the first dipole layer $z = 0$ has only $e^{-2\pi} = 0.0018$ of its amplitude at the start $z = -2c$ of the boundary region. The same ratio of the amplitudes of the inhomogeneous wave holds between the first and second dipole layers.

The rapid decrease of the inhomogeneous waves towards the inside allows one to assume that they are of no importance for the essential process of reflection and refraction. They do not noticeably influence the dynamic condition of the lattice even in the outermost layer of atoms. Besides the attenuation towards the interior, also the amplitude itself of the inhomogeneous waves is of an entirely different order of magnitude than that of the homogeneous waves, especially the homogeneous wave 00 of the potential Π_c . For the term 00 the factor $1/\nu_{lm}$ in Eq. (73) is, because of Eq. (79) of the order of magnitude $1/k$. For the inhomogeneous waves of the higher terms, however, it is only of magnitude 1 Eq. (79a).

From all this it is justifiable to neglect the inhomogeneous waves for the whole inside, including the atoms in the boundary and to take into consideration for the dynamic condition, which must be satisfied at each atom only the potential Π_q and the term 00 of the potential Π_c .

We must briefly inspect this last term. We have

$$\begin{aligned} \nu_{00} &= \sqrt{k_0^2 - \alpha^2 - \beta^2} = \sqrt{k_0^2 - k^2 + \gamma^2} \\ &= k_0 \sqrt{1 - \nu^2 \sin^2 3}, \\ &= k_0 \cos 3_0, \end{aligned} \quad (79b)$$

where $\nu = k/k_0$ is the index of refraction, and (3) and (3_0) are respectively the angles formed with the z axis by the refracted ray (constructed according to the law of refraction) and by the incident ray:

$$\left(\frac{\sin 3_0}{\sin 3} = \nu \right).$$

The only remaining term from Π_c is, when written out in full:

$$\Pi_c^{00} = \frac{i\pi}{2ah} e^{i\alpha x + i\beta y + ik_0 \cos \theta_0 z} \frac{1}{1 - e^{-2ic(k_0 \cos \theta_0 - \gamma)}} \cdot \frac{1}{k_0 \cos \theta_0} \quad (80)$$

The wave normal thus lies in the plane formed by the refracted wave and the z axis (plane of incidence). The phase velocity of the wave will be recognized from

$$\sqrt{\alpha^2 + \beta^2 + \nu_{00}^2} = k_0 \quad [\text{See Eq. (79b)}]$$

as the light velocity in free space. Finally the angle to the z axis is the same angle (θ_0), under which an incident wave would produce the internal wave by refraction.

In the approximation which neglects terms of higher order in ck_0 the factor

$$\frac{1}{1 - e^{-2ic(k_0 \cos \theta_0 - \gamma)}}$$

may be replaced by

$$\begin{aligned} & \frac{i}{2ic} \frac{1}{k_0 \cos \theta_0 - \gamma} \\ &= \frac{1}{2ic} \frac{1}{k_0 \cos \theta_0 - k \cos \theta} = \frac{1}{2ic} \cdot \frac{1}{k_0} \frac{\sin \theta}{\sin(\theta - \theta_0)} \end{aligned} \quad (80a)$$

3.1.2 THE EXTERNAL WAVES Π_a .

The potential Eq. (71) in outer space is very similar to the potential Π_c just discussed. Also here the term 00 exceeds the amplitude and significance of the other terms and one can restrict oneself to that one term.

The term is here

$$H_a^{00} = \frac{i\pi}{2ab} e^{i\alpha x + i\beta y - ik_0 \cos 3_0 z} \frac{1}{1 - e^{2i\pi(k_0 \cos 3_0 + \gamma)}} \frac{1}{k_0 \cos 3_0} \quad (81)$$

The wave thus again proceeds in the 'plane of incidence'. It is the mirror image of the wave Π_c with reference to the surface and has therefore the direction of the reflected wave, which by Snell's construction, is associated with Π_q as the refracted wave.

Its amplitude is different from the inner boundary wave Π_c , since here the factor next to the last is replaced by

$$-\frac{1}{2ic(k_0 \cos 3_0 + \gamma)} = -\frac{1}{2ic} \cdot \frac{1}{k_0} \frac{\sin 3}{\sin(3 + 3_0)} \quad (81a)$$

Figure 8 shows the three waves with their propagation directions and velocities.

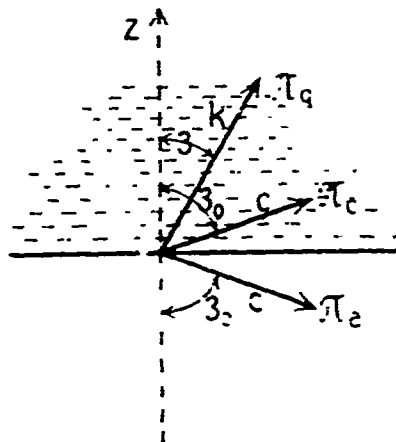


Figure 8

4. FULFILLMENT OF THE DYNAMIC CONDITIONS, THE FRESNEL FORMULAS

The theory of the unbounded crystal has as its objective to show that the assumption of planes of equal dipole phase leads to a dynamically possible state. As a condition of this self-consistency it was found that the propagation velocity of dipole phase and with it that of its accompanying light wave must have a value determined from the dispersion formula.

For the bounded crystal the field by which a dipole is excited results from two potentials, the potentials Π_q and Π_c , from which however there is to be subtracted the part which is emitted by the excited dipole itself. Since Π_q everywhere, even for the outermost boundary atoms, is identical with the potential in the infinite crystal, the same field of excitation, and therefore the same mode of dipole oscillation is generated in the bounded as in the unbounded crystal only if Π_c can be suppressed. As the dynamic condition for the feasibility of the assumption of planes of equal dipole phase we retain here as there only: that one must choose the correct phase velocity q . In the optical case the potential Π_c reduces to a single homogeneous wave that geometrically coincides with the incident wave extended into the inside of the crystal. Therefore the amplitude and phase of the dipole wave can easily be so adjusted, that Π_c and the incident wave cancel.

In order to accomplish the transition from the potential to the field strengths, we take some considerations beforehand which will be systematically discussed in Section 5. In doing so we limit ourselves here to the case of a cubic crystal ($a = b = c$, wave velocity in all directions of equal value, dipole oscillation at right angles to the propagation direction s ; compare dispersion theory). All of the potentials when completely written out have the form

$$ea e^{i\mathbf{x}\cdot\mathbf{s}},$$

where $\mathbf{x}$ is either $\mathbf{k}$ or $\mathbf{k}_0$ and $ea = p$ signifies the moment of the dipole. According to Part I, Section 2, Eq. (4), the field strength is obtained from

the potential by the operation curl curl. We assert here and prove it in Section 5, that this operation in the general case yields the dielectric displacement $\mathfrak{D}$. Only in the case of empty space, $\epsilon = 1$, does it likewise yield the electric field strength $\mathfrak{E}$, which then is equal to $\mathfrak{D}$.

Take as the propagation direction the x axis; we obtain from the above potential

$$\begin{aligned}\mathfrak{D}_x &= \frac{\partial^2}{\partial x^2} a_x e^{ikx} - \frac{\partial^2}{\partial x^2} a_x e^{ikx} = 0, \\ \mathfrak{D}_y &= \frac{\partial}{\partial y} a_x \frac{\partial}{\partial x} e^{ikx} - a_y \frac{\partial^2}{\partial x^2} e^{ikx} = x^2 a_y e^{ikx}, \\ \mathfrak{D}_z &= \frac{\partial}{\partial z} a_x \frac{\partial}{\partial x} e^{ikx} - a_z \frac{\partial^2}{\partial x^2} e^{ikx} = x^2 a_z e^{ikx}.\end{aligned}$$

This means generally expressed: In converting to the field, only the component of the amplitude at right angles to the wave normal produces a contribution. The field of the vector $\mathfrak{D}$ differs from the field of the transverse part of the vector potential by the additional amplitude factor x^2 ($= k^2$ or k_0^2). We have thus to distinguish two cases: one in which the dipole oscillations occur at right angles to the plane of incidence, and the second in which the oscillations occur in the plane of incidence. In the first case the full amplitude a is valid for all three waves; in the second case the amplitude is reduced by a cosine factor.

From the potentials Eqs. (77), (80), and (81) we obtain according to this, the following field strengths for the refracted wave, the boundary wave, and the reflected wave when the oscillations occur in the plane of incidence:

I. Refracted wave:

$$\mathfrak{D}_y = -a \frac{4\pi}{s} \frac{k^2}{k_0^2 - k^2} e^{ikx}; \quad (82)$$

II. Boundary wave:

$$\mathcal{D}_c = a \cos(3-3_0) \frac{2\pi}{r} \frac{\sin 3}{\sin(3-3_0)} \cdot \frac{1}{\cos 3_0} e^{ik_0 r}; \quad (82)$$

(Contd.)

III. Reflected wave:

$$\mathcal{D}_s = a \cos(3+3_0) \frac{2\pi}{r} \frac{\sin 3}{\sin(3+3_0)} \cdot \frac{1}{\cos 3_0} e^{ik_0 r}.$$

Here $v = 8a^3$ is the volume of the elementary cube. If the oscillations are at right angles to the plane of incidence, the first cosine factor is omitted from $\mathcal{D}_c$ and $\mathcal{D}_s$.

Now we relate the incident wave $\mathcal{D}_0$ and the boundary wave according to the condition of cancellation

$$\mathcal{D}_c = -\mathcal{D}_s$$

For oscillations in the plane of incidence, the amplitude of the refracted wave is then expressed by the amplitude of the incident wave as:

$$\begin{aligned} \frac{|\mathcal{D}_r|}{|\mathcal{D}_0|} &= 2 \cdot \frac{k^2}{k_0^2 - k^2} \frac{\operatorname{tg}(3-3_0)}{\sin 3} \cdot \cos 3_0 \\ &= 2 \frac{v^2}{1-v^2} \frac{\operatorname{tg}(3-3_0)}{\sin 3} \cdot \cos 3_0 \\ &= 2 \cdot v^2 \cdot \cos 3_0 \frac{\sin 3}{\sin(3+3_0) \cdot \cos(3-3_0)} \end{aligned} \quad \begin{matrix} * \\ (83a) \\ ** \end{matrix}$$

* The abbreviation tg has not been changed in these equations — it is obviously $\tan$.

** The following formulae are useful for the conversion:

$$\begin{aligned} 1 - v^2 &= \cos^2 3_0 - v^2 \cos^2 3 \\ &= \frac{\sin(3+3_0) \sin(3-3_0)}{\sin^2 3}, \end{aligned}$$

Further:
$$\frac{\cos 3_0 - v \cos 3}{\cos 3_0 + v \cos 3} = \frac{\sin(3-3_0)}{\sin(3+3_0)}.$$

and the amplitude of the reflected wave as

$$\frac{|\mathcal{D}_r|}{|\mathcal{D}_i|} = - \frac{\cos \mathcal{S} + \mathcal{S}_0}{\cos (\mathcal{S} - \mathcal{S}_0)} \frac{\sin (\mathcal{S} - \mathcal{S}_0)}{\sin (\mathcal{S} + \mathcal{S}_0)} = \frac{\tan (\mathcal{S}_0 - \mathcal{S})}{\tan (\mathcal{S} + \mathcal{S}_0)}. \quad *, ** \quad (83b)$$

Finally, for the case in which the oscillations are polarized at right angles to the plane of incidence, the refracted wave would be multiplied by the factor $\cos (\mathcal{S} - \mathcal{S}_0)$ and the reflected wave would be multiplied by

$$\frac{\cos (\mathcal{S} - \mathcal{S}_0)}{\cos (\mathcal{S} + \mathcal{S}_0)}$$

The formulas for oscillation at right angles to the plane of incidence are thus:

$$\frac{\mathcal{D}_r}{\mathcal{D}_i} = 2 \cdot \frac{r^2}{1 - r^2} \frac{\sin (\mathcal{S} - \mathcal{S}_0)}{\sin \mathcal{S}} \cdot \cos \mathcal{S}_0 = 2 r^2 \cos \mathcal{S}_0 \frac{\sin \mathcal{S}}{\sin (\mathcal{S} + \mathcal{S}_0)} \quad (84a)$$

$$\frac{\mathcal{D}_r}{\mathcal{D}_i} = \frac{\sin (\mathcal{S} - \mathcal{S}_0)}{\sin (\mathcal{S} + \mathcal{S}_0)}. \quad (84b)$$

These formulas are in agreement with the Fresnel formulas for the vector $\mathcal{D}$, the dielectric displacement (compare the formula for $\frac{\mathcal{D}_r}{\mathcal{D}_i}$ in Drude Lehrb. d. Optik, 1. Edition. p. 258-262).

The preceding treatment of the problems of reflection and refraction has been developed strictly within the framework of the theory of electrons. It confirms the expectations of the theory of dispersion regarding the influence of the boundary. The fact that an internal boundary wave appears in the electron theory is the point which allows a deeper insight into the problem than the derivation according to Maxwell on the basis of boundary conditions. Even though these boundary conditions were established from the

*, **See footnote on page 68.

concepts of electron theory, the very characteristic property of the boundary lies concealed in the boundary conditions and it does not appear as distinctly as above: why the electrons are bound to oscillate in the described manner in order that they harmonize with the total field of excitation.

The difference of the Fresnel formulas in the two cases of polarization is clarified here in a very direct way as being due to the component formation of the dipole amplitudes. A particular case of this is the familiar explanation of the Brewster angle, as that angle of incidence for which the oscillations of electrons in the body (so far as they occur in the plane of incidence) have the direction of the reflected ray and are thus optically ineffective in producing its amplitude.

5. OBSERVATIONS ON THE TRANSITION FROM THE MICROSCOPIC FIELD TO THE VECTORS $\mathcal{E}$ AND $\mathcal{D}$ OF THE MAXWELL THEORY

The Fresnel formulas were derived in the preceding paragraphs by replacing the potential Π_q by its average value over an elementary parallelepiped; by further performing the operation curl curl on this average value, and by claiming that the macroscopic field of the vector $\mathcal{D}$ may be obtained by this method from the microscopic field.

The reason for this procedure is analyzed in the following.

The vector potential was chosen in view of an easy description of the field of a dipole. In so doing, in particular by setting $\mathcal{E} = \text{curl curl } \mathfrak{P}$ (Part I, Section 2, Eq. (4)), it was assumed that the field propagates in empty space ($\epsilon = 1$). The potential $e^{ik_0 R}/R$ satisfies the differential equation for vacuum

$$\Delta \mathfrak{P} + k_0^2 \mathfrak{P} = 0. \quad (85a)$$

By this rule also in Part I, Section 7, the excitation field was obtained from the exciting potential.

By averaging the total potential $\mathbf{a}''$ to smooth the fast fluctuations, we find that the optical potential now satisfies the equation

$$\Delta \mathbf{\bar{\Phi}} + k^2 \mathbf{\bar{\Phi}} = 0 \quad (85b)$$

k^2 is the constant occurring for the propagation in a medium with the dielectric constant $\epsilon (= v^2)$. But whenever a vector potential satisfies this equation, then the operation curl curl produces from it the vector $\mathbf{\bar{D}}$, the dielectric displacement.

This follows from the definition of the vector potential. It is introduced for a body with the dielectric constant ϵ so as to lead to a solution of Maxwell's equations

$$1/\epsilon \dot{\mathbf{\bar{D}}} = -\text{curl } \mathbf{\bar{E}}; \quad (86a)$$

$$1/\epsilon \dot{\mathbf{\bar{D}}} = \text{curl } \mathbf{\bar{H}}; \quad (86b)$$

$$\text{div } \mathbf{\bar{H}} = 0 \quad (86c)$$

$$\text{div } \mathbf{\bar{D}} = 0 \quad (86d)$$

$$\mathbf{\bar{D}} = \epsilon \mathbf{\bar{E}} \quad (86e)$$

One determines in view of Eq. (86c) curl $\mathbf{\bar{H}}$ such that

$$1/\epsilon \text{curl } \mathbf{\bar{H}} = \mathbf{\bar{D}}$$

the second equation then makes

$$\mathbf{\bar{D}} = \text{curl curl } \mathbf{\bar{H}},$$

whereby also Eq. (86d) is satisfied.

One can now dispose of div $\mathbf{\bar{H}}$ by putting

$$\text{div } \mathbf{\bar{H}} = 0,$$

thus Eq. (86a) in connection with Eq. (86e) requires

$$\Delta \mathfrak{A} - \frac{\epsilon}{c} \mathfrak{A} = 0 \quad \text{or} \quad \Delta \mathfrak{A} + k^2 \mathfrak{A} = 0. \quad (85b')$$

Let it be noted immediately that for an orthorhombic crystal the relationship between $\mathfrak{E}$ and $\mathfrak{D}$ is not given by Eq. (86e) but by the three equations

$$\mathfrak{D}_x = \epsilon_x \mathfrak{E}_x; \quad \mathfrak{D}_y = \epsilon_y \mathfrak{E}_y; \quad \mathfrak{D}_z = \epsilon_z \mathfrak{E}_z, \quad (86e')$$

and that thus the three components of $\mathfrak{A}$ satisfy different equations Eq. (85b), in which, instead of the uniform ϵ are substituted $\epsilon_x, \epsilon_y, \epsilon_z$ for the three components $\mathfrak{A}_x, \mathfrak{A}_y, \mathfrak{A}_z$. But here as in the isotropic body (cubic crystal) it is always true that

$$\text{curl curl } \mathfrak{A} = \mathfrak{D}. \quad (85c)$$

So much for the reasons that the method used provides the Fresnel formulas for the vector $\mathfrak{D}$ and not for $\mathfrak{E}$. However, a remark may be added on the transition from the differential equations Eqs. (85a) to (85b), which is connected with the definition of the values $\mathfrak{E}$ and $\mathfrak{D}$ and which throws an interesting light on the logical aspect of the first part of the theory of dispersion.

The total potential satisfies Eq. (85a) before averaging; Eq. (85b) after averaging. The integration over the elementary parallelepiped can produce a change in the differential equation only because it includes an integration over the pole. For the integration of $\mathfrak{P}$, which would become infinite as $1/R$, the pole is insignificant. However, if $\Delta \mathfrak{P}$ is integrated over the pole, the near vicinity of the pole contributes, as is well known, a term $4\pi p$, where $p = e \cdot a$ signifies the moment of the dipole, while the remainder of the integration space produces $\Delta \mathfrak{P}$ where $\mathfrak{P}$ denotes the average potential. The average of the Eq. (85a) over the volume v thus gives:

$$\Delta \bar{\Phi} + k_0^2 \bar{\Phi} + \frac{4\pi}{v} \bar{p} = 0. \quad (85d)$$

If now the average of the potential is the potential of the Maxwellian quantity $\bar{\Phi}$ then it must also satisfy the propagation equation for $\bar{\Phi}$. That is, we must also have

$$\Delta \bar{\Phi} + k^2 \bar{\Phi} = 0. \quad (85e)$$

From these two equations, the average value of the potential itself follows:

$$(k_0^2 - k^2) \bar{\Phi} = -\frac{4\pi}{v} \bar{p} = -4\pi \bar{p}$$

or

$$\bar{\Phi} = -\frac{4\pi}{k_0^2 - k^2} \bar{p}. \quad (85f)$$

But this is exactly the result from the theory of dispersion [Part I, Section 6, Eq. (31)] which was again derived in Part II by averaging over the sum of the dipole fields [Eq. (77)].

One sees: If it is assumed that the average value of the potential is the potential of the optical field and therefore satisfies Eq. (85e), then it is not necessary first to calculate the exact fields produced by the dipole oscillations. Rather, one can infer the Maxwell field quantities directly from the assumption made concerning the dipole oscillations. Whether the assumed state of dipole oscillations is dynamically possible, one is, however, not in a position to decide without going into the details of the field. But only a knowledge of the fields in the immediate vicinity of the dipoles (the excitation fields) would be required for finding the answer.

The logical place of our computation of the total field from the assumed dipole oscillations is, in Part I, the theory of dispersion, an existence proof. Without using the assumptions necessary for the derivation of Eq. (85f), it is shown that there exists an average value of the

potential which satisfies Eq. (85b) and is therefore the potential of the optical field. In the theory of reflection and refraction on the other hand the computation of the total field serves the purpose of resolving the dynamics of the problem. It could therefore not be omitted, even if relationship Eq. (85f) were used.

In conclusion we compare our definition with that used by H. A. Lorentz.⁽¹⁾ The latter is, written in our units

$$\mathfrak{D} = \mathfrak{E} + 4\pi \bar{\mathfrak{p}},$$

or

$$\mathfrak{D} = \frac{\epsilon}{\epsilon - 1} \cdot 4\pi \bar{\mathfrak{p}} = -4\pi \cdot \frac{k^2}{k_0^2 - k^2} \bar{\mathfrak{p}}. \quad (85g)$$

The same relationship between the dielectric displacement and polarization we also obtain from the relationships Eqs. (85c) and (85f) in all cases in which the assumption of the average polarization $\bar{\mathfrak{p}}$ is descriptive of a state which is consistent with the Maxwell Eqs. (86a) to (86d), that is, whenever

$$\Delta \bar{\mathfrak{p}} + k^2 \bar{\mathfrak{p}} = 0; \quad \text{div } \bar{\mathfrak{p}} = 0$$

We have then from the definition Eq. (85c)

$$\begin{aligned} \mathfrak{D} &= \text{curl curl } \bar{\mathfrak{p}} = -4\pi \frac{1}{k_0^2 - k^2} \text{curl curl } \bar{\mathfrak{p}} \\ &= -4\pi \frac{k^2}{k_0^2 - k^2} \bar{\mathfrak{p}} \end{aligned}$$

in agreement with Lorentz.

(1) Enzyklop, d. math. Wiss. V. 14 (XXI), (XXV), (XXXI) and No. 30.

The definition contained in Eq. (85c) thus leads to the familiar vector of dielectric excitation or displacement for all states of dipole motion which permit comparison with a Maxwellian field.

6. WAVE EXIT, TOTAL REFLECTION

In order to study the exit of a plane wave from the body, which as before fills the upper half space, only small changes in the final formulas are required. This case differs from the preceding one only by the reversal of the wave direction, that is, the values α, β, γ have to be replaced by their negative values. We make this change, and the principal terms of the three potentials Eqs. (77), (80) and (81) then read:

$$\Pi_a = \frac{i\pi}{2ab} e^{-i\alpha x - i\beta y - i\nu_{00}z} \frac{1}{1 - e^{2ic(\nu_{00} - \gamma)}} \cdot \frac{1}{\nu_{00}}, \quad (87a)$$

$$\Pi_c = \frac{i\pi}{2ab} e^{-i\alpha x - i\beta y + i\nu_{00}z} \frac{1}{1 - e^{-2ic(\nu_{00} + \gamma)}} \cdot \frac{1}{\nu_{00}}, \quad (87b)$$

$$\Pi_\gamma = \frac{\pi}{2abc} e^{-i\alpha x - i\beta y - i\gamma z} \frac{1}{k_0^2 - k^2}. \quad (87c)$$

Here ν_{00} has the old value:

$$\nu_{00} = \sqrt{k_0^2 - \alpha^2 - \beta^2} = k_0 \sqrt{1 - \nu^2 \sin^2 \theta}. \quad (87d)$$

In Π_q we recognize again the old wave of velocity q , now, however, running towards the boundary. Its amplitude is unchanged.

Π_a as well as Π_c have the phase velocity c and have their angle with reference to the z axis preserved, but the angles to the x and y axes

changed, so that the negative axis directions replace the positive ones. Also the previously given amplitudes are changed. We substitute as above [Section 3 Eq. (80a)]:

$$\frac{1}{1 - e^{-\frac{1}{2}ic(v_{00} - \gamma')}} \quad \text{by} \quad -\frac{1}{2ic} \frac{1}{v_{00} - \gamma'}, \quad (87')$$

$$\frac{1}{1 - e^{-\frac{1}{2}ic(v_{00} + \gamma')}} \quad \text{by} \quad \frac{1}{2ic} \frac{1}{v_{00} + \gamma'}. \quad (87'')$$

The pictures of the waves which result from incidence and exit are compared with each other in Figure 9.

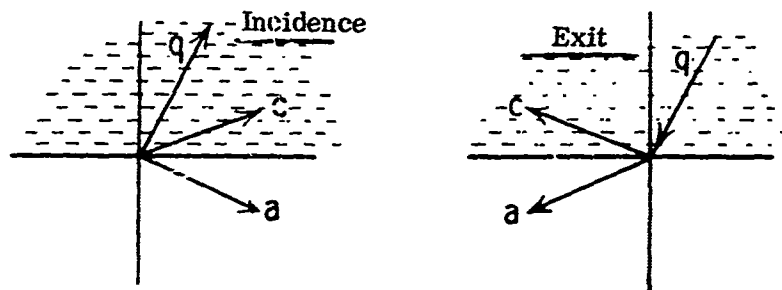


Figure 9

One now sees easily how to fulfill the dynamic requirements. The internal wave of velocity c has to be removed, since it would disturb the electrons whose oscillations are balanced only against the wave of velocity q . This is accomplished by considering the dipole oscillations coupled to the internally reflected wave q' which would produce an internal boundary wave c' of the same direction as c . (Figure 10). Its amplitude is determined by the requirement that c and c' cancel each other. This gives at

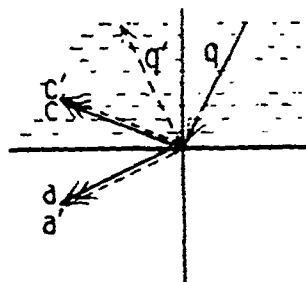


Figure 10

the same time, the condition for the newly introduced dipole oscillations to harmonize with their field q' . Finally, the field which appears in the outer medium as the transmitted wave of amplitude a'' arises from both component waves q and q' that is, from a and a' superimposed. We set the amplitude of the wave striking the boundary surface from inside equal to 1. Thus we have the following formulas for the amplitude values of the potential:

$$\begin{aligned} q &\dots 1 \\ c &\dots -\frac{1}{2} \frac{1}{v_{\infty} + \gamma} \cdot \frac{1}{v_{\infty}} \cdot (k_0^2 - k^2) \quad \text{from Eqs. (87b) and (87'')} \end{aligned} \quad (88a)$$

Therefore

$$c' \dots + \frac{1}{2} \frac{1}{v_{\infty} + \gamma} \cdot \frac{1}{v_{\infty}} \cdot (k_0^2 - k^2)$$

and further from Eqs. (77), (80), and (80a)

$$\begin{aligned} q' &\dots -2 \cdot \frac{(v_{\infty} - \gamma) \cdot v_{\infty}}{k_0^2 - k^2} |c'| \\ &= -\frac{v_{\infty} - \gamma}{v_{\infty} + \gamma}, \end{aligned} \quad (88b)$$

and hence from Eqs. (87a) and (87'), and Eqs. (77), (81), and (81a)

$$a'' = a + a' \dots$$

$$\frac{1}{2} \frac{1}{r_{00} - \gamma} \cdot \frac{1}{r_{00}} (k_0^2 - k^2) - \frac{1}{2} \frac{r_{00} - \gamma}{(r_{00} + \gamma)^2} \frac{1}{r_{00}} (k_0^2 - k^2). \quad (88c)$$

In going over to the fields, q and q' have to be multiplied with the factor k^2 , the wave a'' in outside space however by k_0^2 . If we continue to set the amplitude of the field strength of the wave q equal to 1, then the amplitude a'' contains the factor $k_0^2/k^2 = 1/\nu^2$. By utilizing the formulas mentioned in the footnote section 4, page 68, the amplitudes of the internally reflected and the transmitted wave, for the case that the oscillation is at right angles to the plane of incidence, become

$$q' \dots - \frac{\sin(3 - 3_0)}{\sin(3 + 3_0)}; \quad a'' \dots \frac{2}{\nu^2} \frac{\cos(3 \sin 3_0)}{\sin(3 + 3_0)}. \quad (89)$$

These are the same values as are obtained from the Fresnel formulas Eqs. (84a) and (84b) when the angles (3) and (3_0) are interchanged and ν is replaced by $1/\nu$ corresponding to the progression from an optically denser to an optically thinner medium.

Fulfillment of the dynamic conditions has thus provided the correct amplitudes also in the case of wave exit. It was tacitly assumed for the last conversions that

$$r_{00} = \sqrt{k_0^2 - \alpha^2 - \beta^2} = k_0 \sqrt{1 - \nu^2 \sin^2 3}$$

may be replaced by $k_0 \cos 3_0$. This implies that we are not dealing with the case of total reflection.

Let us assume now, that the angle of incidence (3) of the internal wave with the boundary surface is so large that total reflection occurs. ν_{00} is then imaginary and the two boundary waves associates with the q

wave in inner and outer space are no longer homogeneous, but inhomogeneous. They agree in this regard with the higher terms of the series Eqs. (71) and (73) for the potentials Π_a and Π_c whose principal (00) terms they were. But compared with these higher terms, for instance the term (01), they show a quantitative difference in the degree of inhomogeneity.

While

$$r_{00} = \sqrt{k_0^2 - \alpha^2 - (\pi/b - \beta)^2} \simeq i\pi/b$$

has an imaginary value of the order of magnitude of $1/b$,

$$r_{00} = ik_0 \sqrt{r^2 \sin^2 3 - 1} \simeq ik_0$$

is only of the order of magnitude k_0 , that is, $1/\lambda$. The inhomogeneous wave (00) thus reaches with noticeable strength as far down in depth as the value of the wavelength, while the higher order waves do not reach beyond the first atomic layer. Also, as before, since the factor $1/\nu_{lm}$ appears in the absolute value of the amplitude, the higher inhomogeneous waves are much weaker than the principal wave (00). Physically speaking, the higher inhomogeneous terms describe the perturbation at the surface, which is caused by the discontinuous nature of the surface. The (00) wave, on the other hand, is produced by the events in the interior and is independent of the smallness or coarseness of the surface lattice of atoms.

The dynamic conditions remain the same for the total reflection as for the previously discussed exit of a wave from the medium. The internal inhomogeneous wave c must be removed by considering the internally reflected wave. Inside the medium, the two inhomogeneous waves c and c' are superimposed. For the reflected wave in the interior we obtain as above Eq. (88b) the amplitude

$$-\frac{r_{00} - r}{r_{00} + r} = \frac{k \cos 3 - ik_0 \sqrt{r^2 \sin^2 3 - 1}}{k \cos 3 + ik_0 \sqrt{r^2 \sin^2 3 - 1}}.$$

Its absolute value, which is the important quantity, is 1 since numerator and denominator are complex conjugates.

The potential of the inhomogeneous wave in outer space has at the boundary the amplitude [compare Eq. (88c)]

$$\frac{1}{r_{\infty}} (k_0^2 - k^2) \left(\frac{1}{r_{\infty} - r} - \frac{r_{\infty} - r}{(r_{\infty} + r^2)} \right) = \frac{2 \cos \theta}{(i \sqrt{r^2 \sin^2 \theta - 1} + r \cos \theta)}.$$

The absolute value of this is

$$\frac{2 \cos \theta}{\sqrt{r^2 \sin^2 \theta - 1}}.$$

For the field of the electric vector, we have to add the factor $k_0^2/k^2 = 1/v^2$ for the same reason as previously. That the result coincides with the one obtained from Maxwell's theory, follows from the fact that it can be obtained here as above from the Fresnel formulas by replacing $\cos \theta_2$ by $i \sqrt{r^2 \sin^2 \theta - 1}$.

7. TRANSITION TO THE CASE OF X-RAYS *

The series Eqs. (71), (72) and (73) for the field inside and outside the crystal are valid in all strictness on the basis of the assumptions. The discussion of the previous paragraphs indicates that for the case of optical waves ($ak \ll 1$) the three series reduce essentially to their principal terms, in which all indices are zero.

It is otherwise when x-rays penetrate the lattice. ** We consider the field in outer space, whose potential is the series Eq. (71) (The same consideration is associated also with the exit of the x-rays.) There exists

* This case I have already treated on the basis of formula (71) in *Phys. Zeitschr.* 14:465, 1913.

**The lattice is naturally thought of as being perfectly regular. In particular the attenuating thermal motion is completely neglected.

now a whole series of terms for which

$$r_{lm} = \sqrt{k_0^2 - \left(\frac{l\pi}{a} - \alpha\right)^2 - \left(\frac{m\pi}{b} - \beta\right)^2}$$

is real, and which therefore describe homogeneous plane waves. Among these waves those are dominant for which the denominator of Eq. (71) vanishes. The condition for its first factor to vanish is that

$$c(r_{lm} + \gamma) = n\pi$$

where n is a positive or negative integer. Restated, this condition becomes

$$(l\pi/a - \alpha)^2 + (m\pi/b - \beta)^2 + (n\pi/c - \gamma)^2 = k_0^2. \quad (90)$$

Note that here (certainly to a high degree of approximation)

$$\alpha^2 + \beta^2 + \gamma^2 = k_0^2 (= k^2).$$

Equation (90) leads directly to the sphere construction ⁽¹⁾ for the determination of the direction of the interference maxima: It is, in fact, identical with the Lane conditions ⁽²⁾ for the occurrence of an interference maximum.

That the amplitude of the waves appears to be infinitely large is due to the cooperation of the infinite number of dipoles and the neglect of absorption.

Observe moreover, that in the interior of the crystal according to the potential Π_q likewise each of the waves (l, m, n) attains a maximum strength for which Eq. (90) is satisfied. The field in the interior consists essentially of the interference-favored waves. *s one of these, and not further

(1) P. P. Ewald, loc. cit. or Enzyklop. d. math. Wiss., Article Lane, V. 24, No. 48.

(2) M. v. Lane, Ann. d. Phys. 41; 971, and Enzyklop. d. math. Wiss. Article Lane V. 24, No. 46.

distinguished from the others, there exists the (000) wave, which in the optical case was the only dominant one, namely the refracted wave.

One apparently obtains in Eq. (71), an additional mode of infinite amplitude, namely when the second denominator vanishes:

$$\nu_{lm} = 0.$$

It follows, however, from the derivation of formula Eq. (71) that application of Eq. (71) to this case is not permitted, (the pole is in this case of second order).

It is instructive to consider the transition from the case of x-rays to the case of optics. To that effect, we let k gradually decrease -- we can think of the crystal lattice shrinking while we hold the wavelength constant. There are then a diminishing number of index pairs l, m , which leave ν_{lm} real. Keeping l, m constant, ν_{lm} becomes ever smaller. This means that the fan of interference rays which is present for the shorter x-rays, spreads out. A moment comes where the rays which are diffracted the most have $\nu_{lm} = 0$. From there on ν_{lm} for these rays would become and remain imaginary. As soon as the wave directions approach the crystal surface the waves become inhomogeneous. The spreading of the pencil of interference rays can, of course, be understood by the increasing dispersive power of the shrinking lattice which spreads the spectra apart. Only the ζ_0 ray (indicated thick in Figure 11) does not change its position in this

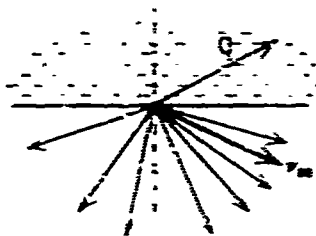


Figure 11

process. It is this ray that finally remains as the only one of the reflected spectra that survives in the optical domain.

Acknowledgments

I would like to express my appreciation to those individuals at AFCRL who assisted in clarifying several points in my translation and particularly to Dr. H. Zschirnt who assisted me on a number of occasions.

Special thanks are expressed to Prof. P. P. Ewald who completely revised or rewrote several portions of my translation so that correct statements were made rather than erroneous ones as well as making numerous smaller corrections. I am also indebted to him for the supplementary material that he has written for this publication. Without his generous assistance and contributions, none of this publication would have been feasible.

L. M. Hollingsworth

Appendix A

Postscript 1970
by P. P. Ewald

DEVELOPMENTS ARISING OUT OF PARTS I AND II
ON THE FOUNDATIONS OF CRYSTAL OPTICS

Before the end of the summer semester 1910 at the University of Munich, I asked my revered teacher, Professor Arnold Sommerfeld, for permission to start on thesis work under his guidance. He suggested for subjects some ten different applications of Maxwell's equations to boundary value problems, such as radio wave propagation, or the self-inductance of coils of various shapes. At the bottom of the list stood the problem: Does anisotropic arrangement of resonators produce double refraction? This was the only problem that ended with a question mark, and the Professor explained that he himself saw no clear way of tackling the problem and that it would be much more risky, from my point of view, of choosing this problem rather than one of those where he could foretell what the solution would require in the way of mathematics. He explained that crystallographers had long since developed elaborate theories of crystal structure but that these remained purely geometrical speculations as long as no quantitative consequences had been drawn -- and that if I could find an answer, it might help to corroborate or to refute the ideas on crystal structure.

Of all the subjects he proposed, only this one evoked any enthusiasm in me, and so, after thinking it over, I returned to him with the decision that I would like to work on it. Professor Sommerfeld gave me the reprint of a paper by M. Planck on the theory of dispersion - and with that I started on a solitary hike up the Rhine valley in Switzerland for the next weeks of summer vacation.

I was not much wiser when I returned to Sommerfeld in the fall. Not only had I not progressed with my problem, but I even found great difficulty in understanding what Planck was doing, and Planck had always been praised to me as a master of clarity. There seemed to me to be some confusion in his calculation of the field created by the resonators (dipoles) in the medium and the "incident field" which seemed to exist in addition.

Professor Sommerfeld, meanwhile, had also given some thought to the problem. Using the Hertz potential, he had tried to sum up the fields of the spherical waves issuing from the dipoles at the nodal points of the orthorhombic lattice and had recognized that the main obstacle in this summation was the fact that the dipole coordinates X, Y, Z , entered in the Hertz potential only in the irrational expression

$$R = \sqrt{(X-x)^2 + (Y-y)^2 + (Z-z)^2}$$

where (x, y, z) denote the field point. He then showed me a representation of $\frac{1}{R} e^{ikR}$ by an integral containing Bessel functions of r , the two-dimensional analogue to R , and an exponential factor in which the exponent was linear in the third coordinate $(Z - z)$. The summation over the third coordinate could then be readily performed under the integral sign, for it was nothing but the sum of a geometrical series. Could we not, Sommerfeld suggested, bring the other two coordinates into a similar linear form? For the next few months, I tried, in vain to achieve this - then, suddenly, I saw the light and obtained the desired representation in the form of the triple integral [Part I, Section 5, Eq. (23)]. When I arranged to show Sommerfeld my recent advance, he greeted me with the words: "I too can show you

something that should be helpful" - and he developed the very same integral which he had derived in a less cumbersome way.

The path now lay open for finding the total potential by means of the residue method and I rapidly arrived at the expression Part I, Section 6, Eq. (39).

My happiness over this first tangible result was not to last long: I found it impossible to subtract from this total potential that part that was contributed by the test dipole. Since the oscillations of the test dipole are caused only by the radiation coming from the other dipoles, this subtraction appeared essential for the determination of the "field of excitation". The impossibility of a direct subtraction comes from the fact that in the expression Part I, Section 6, Eq. (39), all individual dipole fields are jumbled together and recast into a sum of nonlocalized plane waves in which individual contributions can no longer be recognized. (I was not aware, at the time, that what I had obtained was a Fourier development of the total field.)

Success in the removal of the field of the test dipole came through Sommerfeld's assistant, P. Debye. At a skiing holiday in Mittenwald, Easter 1911, Sommerfeld showed him our quandary. With one glance at the expression of the total potential, Debye said: "This is quite simple; you have to use Riemann's method of bringing the denominator of the sum term into the exponent of an exponential function by introducing a new integration. The integrand then becomes a theta-function and you can apply the transformation theorem of the theta functions to it." The whole "consultation" probably lasted no more than a quarter of an hour. Debye was not only an avid reader of classical physics, but also had an extraordinary power of seeing through mathematical formalism. He did not know at the time in which of Riemann's papers this method occurred - nor did I ever find out.

When I had finally understood Debye's advice, I transformed the total potential in the way he had suggested. This is a rather round-about way of obtaining the theta functions and in the publication I followed a more direct way by introducing the simple-integral representation of $\frac{1}{R} e^{ikR}$, [Part I, Section 5, Eq. (18)], which leads directly to theta functions.

For an orthorhombic lattice, the theta function in (X, Y, Z) is the product of three such functions of one variable each, and the transformation is easily written down. The integrands in the two forms thus obtained are equivalent at any point of the integration over ϵ . In fact, one is the Fourier development of the other. It is, therefore, permissible to use one integrand up to an arbitrary point E of the whole integral path, and then change to the other form for the remainder of the integration. In the integral from E to ∞ individual contributions to the potential are still recognizable, and the contribution of the test dipole can be omitted; whereas in the integration from 0 to E , the Fourier-transformed integrand is used in which individual contributions are blotted out. By suitably choosing the transition point E , one arrives at having rapid convergence of the sums in both integrals. A particular advantage lies in the fact that by shifting E , rapid convergence can be shifted from one sum to the other. Since the sum of the two part integrals is independent of E , an easy check on the correctness of the numerical work is afforded by using two different values E .

I have later extended this method of calculating lattice potentials and Madelung constants from orthorhombic to general lattices.* There, the theta function of the three space coordinates no longer splits into the product $\theta(x) \cdot \theta(y) \cdot \theta(z)$. The transformation property of the space-theta was contained in the main text on theta function,** but in a horrible form. By using the concept of the reciprocal lattice, I could re-state it in a way akin to the tools of the mathematical physicist. Since then, the theta function method has been the favored one for calculating lattice energies and potentials. The neat way in which the formal splitting of the integral for the potential produces two rapidly convergent parts was interpreted physically in a later paper of mine.***

* Ann. d. Phys. Lpz. (1921) 64:253-287.

** A. Krazer Lehrbuch der Thetafunktionen, Teubner, Lpz. 1903.

*** Nachr. Ges. d. Wiss. zu Göttingen (1938) Math-phys. Kl. 3:55-64.

For bringing my thesis to a conclusion, two points had to be discussed:

(1) Did the formal expression obtained for the refractive indices of waves travelling along the x, y, and z directions of the orthorhombic lattice guarantee that the optical behavior of the medium would conform to the laws of crystal optics and be valid also for other directions of light propagation? This is answered in the affirmative sense in Part I, Section 10.

(2) Is the birefringence found, a major or negligible effect? It was not easy in 1911 to carry out such a comparison because in spite of the existence of an elaborate geometrical theory of crystal structure the question was left wide open what kind of particle - molecules or groups of molecules - was to be placed in the assigned positions. The birefringence might well have its origin in an inherent anisotropy of these particles rather than in the contribution caused by their anisotropic arrangement. What I needed for the comparison was a crystal of known birefringence which was built according to the simple orthorhombic lattice postulated in my calculations. The famous old professor of mineralogy and crystallography, Paul Groth, to whom I took my question, exclaimed after some hard thinking: "There is one crystal which is most certainly built according to the lattice you want, for it has three marked cleavage planes at right angles to one another. This crystal is anhydrite, CaSO_4 . Try that." So I did. The result of my calculations is to be found in Part I, Section 11. It was shattering my secret hopes of obtaining values showing at least some similarity to the actual birefringence of anhydrite.

So, with some resignation, I stated that the order of magnitude of the birefringence due to the lattice arrangement is the same as the observed one. From this it followed that in any future calculation of birefringences the effect of the anisotropic arrangement would have to be taken into account.

It was only years later, in 1925 and 1926, that the true structure of anhydrite was established as consisting, essentially, of SO_4 tetrahedra arranged according to a b-face centered, not a primitive, orthorhombic lattice.

In my thesis, I made some attempt to discuss what would be implied in an extension of the theory to more complicated assumptions regarding the crystal structure. The two sections, one on the influence of infrared proper vibrations, and the other on the refraction by a general periodic arrangement of dipoles, did not lead to any physical results except showing that some of the conclusions reached under the simpler assumptions would be invalidated; they were, therefore, omitted from the paper published in the Annalen der Physik.

The calculation of double refraction became much more realistic once the determination of crystal structures had become possible after Laue's and Bragg's work in 1912. In particular, E.A. Hylleraas was very successful in accounting for the birefringence of the tetragonal calomel (Hg_2Cl_2)^{*} and of the trigonal Quartz (SiO_2)^{**} by an extension of the theory to more general structures than the simple dipole lattice. In the case of quartz, Hylleraas could also account for the optical rotatory power, using an extension of my theory which had meanwhile been worked out by M. Born in order to derive optical activity from the presence of screw axes in the crystal structures. In all these cases, it was sufficient to assume the atoms to be represented by isotropic dipoles; later, cases became known where an inherent anisotropy of the resonators had to be assumed in addition to the effect of their anisotropic arrangement.^{***}

The theory of "Dispersion and Double Refraction of Electron Lattices (Crystals)," was first printed under this title in Göttingen 1912 and submitted as Ph.D. thesis to the Philosophical Faculty (II. Section) of the University of Munich under the date of 16th February 1912. The present

* Zs. f. Physik 1926, 36:859-896.

** Zs. f. Phys. 1927, 4:871.

*** See, for instance, H. C. Bolton, W. Fawcett, I. D. C. Garney, Proc. Phys. Soc. (1962) 80:199-208 and (1963) 82:33-46.

translation is from a subsequent publication* where, under a new general title Foundations of Crystal Optics, an introduction is given not only to the theory of dispersion, but also to Part II which is called Theory of Reflection and Refraction. Two more papers appeared under the same general title, namely in 1917**, and twenty years later.***

The latter two parts deal with the case that the wavelength of the optical field is shorter than the distance between neighboring atoms in the crystal. The strictly valid transformations of the sum of spherical dipole waves from Part I are again used, but their discussion proceeds along entirely different lines for the X-ray than for the optical case. The same condition of consistency of the dipole oscillations with the field created by them is applied and its result is the determination of the mode of propagation of the optical field in the interior of an unbounded crystal. The X-optical field, however, is more complicated than that of visible light, for it consists not of a single plane wave, but of a pencil of such waves which are coupled together. The dynamical condition for the propagation of the X-optical field is expressed geometrically by a surface of a similar nature to the surface of normals in visible optics (Part I, Section 10); for X-rays this is called the surface of dispersion. Once this surface has been determined, all possible modes of propagation of self-consistent X-fields in the interior of the unbounded crystal are known. There remains, however, the same problem as in the case of visible light, namely of relating the internal optical field to the external one for a crystal with boundary, i. e., a crystal filling only one half-space. This problem was treated for visible light in

* Annalen der Physik (Lpz) (1916) 49:1-38.

** Part III: The Crystal Optics of X-Rays, Ann. d. Phys. Lpz (1917) 54:519-597.

***Part IV: Establishment of a General Condition of Dispersion in Particular for X-Ray Fields, Zeitsch. f. Kristallographie (A) (1937) 97:1-27.

Part II of the series, and the same principle, established there, could later be applied to answer the X-ray case.

The so-called Dynamical Theory of X-ray Diffraction of Part III was restricted by the assumption that the crystal could be represented by a simple ('primitive') lattice of dipoles. The restriction to an orthorhombic lattice was lifted because one had meanwhile learned to handle nonorthogonal lattices by means of the concept of reciprocal vector sets.* In the simple case of only two plane waves forming the X-optical field, α , could deal with a general crystal structure, thus getting rid of the first restriction (on the crystal) by introducing a new one (on the field). The resulting intensity formula (but not the entire theory) was published in a paper of mine.**

This, however, left the general case unanswered in which an X-optical field consisting of n component plane waves travels through a general periodic crystal structure. This problem was tackled in Part IV and brought to a general, if formal, solution. The ideal aim of this theory was to obtain a uniform treatment valid for short X-wavelengths as well as for the long ones of visible light and beyond this for the static dielectric constant, and all that in a quite general periodic medium. The formal answer obtained in this paper deals only with the internal condition of self-consistency which, however, is the clue to the boundary problem. An application of this general theory to the propagation of an X-field consisting of three plane waves in a diamond-type structure is to be found in the paper by Ewald and Hénoc.***

The dynamical theory of X-ray diffraction was recast in a different form in 1931 by M. von Laue. The crystal was considered to be a continuous periodic medium in which a periodic dielectric constant affects the propagation of the X-rays. Laue's theory thus departs from the strict

* Introduced by Willard Gibbs in his lectures 1881-1884, but first published in Gibbs-Wilson Vector Analysis, 1901.

** Physik. Zeitsch (1925) 26:29-32, and 1926, 27:182.

*** Acta Crystallographica 1968, 24:5-15.

concepts of the Lorentz electronic theory of matter and makes use of the Maxwellian material constants and field quantities and their relations. A justification for this step was given by M. Kohler in 1935 on the basis of wave mechanics and perturbation theory. The chief difference between Laue's and my own version of the theory is that Laue makes use of Maxwellian type boundary conditions at the surface of a half-crystal in order to connect the external and the internal field, whereas in my theory this is achieved by the condition of self-consistency which has to be satisfied at each dipole throughout the depth of the crystal. The results are the same in both versions.

The dynamical theory of X-ray diffraction has achieved increasing importance where crystals are used as optical instruments of considerable perfection. Of greater practical importance, however, is the application of the concepts created in this theory to the case of electron diffraction. The scattering effect of matter is much stronger for electrons than for X-rays, and therefore the multiple scattering effects are more prominent with electrons. Since these are included in the "dynamical" theory, this theory has to be applied to electron diffraction at an earlier stage of refinement than is necessary in X-ray diffraction. The theory of electron diffraction was first given by H. A. Bethe in 1928, and it uses many of the concepts introduced for X-rays. It is a scalar theory of diffraction dealing with Schrödinger's ψ function as the field quantity rather than with the vectorial field strengths of the optical theory. The theory is an unfailing guide in the discussion of the highly complex observations of electron diffraction.

With the advent of neutron diffraction, the concepts and methods of the dynamical theory found an important new field of application.*

After this short survey of the wide field of developments beginning with the two parts of the series which are to be found here in translation, let us return to Part II with some comments. It contains the connection between the internal and the external field in a "half-crystal," i.e., a lattice filling

* See for instance C. G. Shull (1968) Phys. Rev. Letters, 21:1585-1589.

one half space and cut off at the surface $z = 0$. The half-crystal is still infinite, but it is bounded. Strictly speaking, it has no boundary in the sense of the Maxwellian theory for a continuous medium, and therefore the theory does not contain boundary conditions which would be valid over the entire plane $z = 0$. Instead, the half-crystal is an open structure with empty space not only outside the lattice, but also inside between the dipoles. The peculiar boundary conditions for a continuous medium, like E_{tang} and D_{norm} the same on both sides of the boundary, are replaced in the open structure by the condition of self-consistency at the site of each dipole, no matter at what depth under the surface it lies.

It was already pointed out in Part I, that in the theory of dispersion there was no room for an incident wave - in fact where should it come from if the crystal fills all space? In Part II, it is shown that by restricting the summation of spherical dipole wavelets to sources on one side of the plane $z = 0$, extra fields are created which travel on both sides away from this plane with the velocity c , or in the higher terms of the series, are inhomogeneous surface waves; i.e., waves whose (small) phase velocity is directed along the surface while their amplitude decreases exponentially with increasing distance from the surface. The term 'boundary waves' is used for these waves in order to encompass both the homogeneous and inhomogeneous ones. In other places* I have called them epiwaves, in contrast to the field in the unbounded crystal which consists of mesowaves.

In Part II, the boundary waves arise from residue formation at special series of poles in the plane of integration. A less formal understanding of their occurrence is desirable and can be obtained in two different ways.

The first is a simple application of the sum of a geometrical series, say of quotient Q . The infinite series, $1 + Q + Q^2 + \dots$ has the sum $1/(1-Q)$; the finite series, $1 + Q + Q^2 + \dots + Q^{(p-1)}$ has the sum $(1 - Q^p)/(1 - Q)$, that is, an additional term is caused by the termination of the series.

* For instance in Gomer and Smith, Structure and Properties of Solid Surfaces, Univ. of Chicago Press 1953, Ch. II.

Now we consider the spherical waves issuing from a plane of equal dipole phase. These waves will, soon after leaving that plane, merge to form a plane wave travelling with the velocity c away from the plane of the dipoles. At a field point z there arrive such plane waves emitted from the nearest, the next nearest and successively more distant planes of dipoles. Since the assumed velocity of the dipole phase is q , there will be at z a time lag ϕ between contributions from neighboring dipole planes equal to $a(1/q - 1/c)$ where a is the distance between the dipole planes. We thus obtain at z a sum of waves of equal amplitudes but with a phase factor $Q = e^{j\nu\phi}$ between them where ν is the frequency and $j = 2\pi\sqrt{-1}$. The resulting amplitude at z is the sum of the geometrical series

$$\sum_n e^{jn\nu\phi} = \sum_n Q^n .$$

If the series is continued to infinite values of n (assuming a slight absorption for the sake of convergence), the resulting amplitude is $1/(1-Q)$. If we have to terminate the summation at the $(N-1)^{st}$ layer above z , an additional term $Q^N/(1-Q)$ arises. This term corresponds to the boundary wave. A full discussion of the theory along these simplified lines is the subject of a paper of mine.*

The second, even more physical way of understanding the field in the interior of the half-crystal is this. Imagine first all space filled with the dipole lattice oscillating as assumed in Part I. The result is the field derivable from the potential Part I, Section 6, Eq. (30) which we now call for short the mesofield. Next cancel the contributions of the dipoles above the plane $z=0$ by adding in the upper half of space fictitious dipoles of the opposite phase. Then the total field in the lower half-space is the sum of the mesofield and the field produced in the lower half-space by the fictitious

* Fortschritte der Chemie, Physik und physikal. Chemie Serie B (1925) 18, Heft 8 (Berlin, Gebr. Bornträger Verlag). See also Physica (1924) 4:234-251.

dipoles, i. e., outside the fictitious medium. This is, of course, a field consisting of waves of velocity c ; in fact, it is the field of the internal boundary waves of the half crystal. The fictitious dipoles, apart from destroying the mesofield in the upper half of space, also create there an internal boundary wave in their (fictitious) medium. For the real half-crystal this is its external boundary wave which then becomes the reflected wave when the relationship to the wave incident on the half-crystal is established.

This physical explanation is not too different from the foregoing explanation by means of the termination term of the finite geometrical series. For the formula for the finite sum can be interpreted as the instruction: sum first the infinite series and then subtract all terms beyond the point of termination ($p - 1$); i. e., the same infinite series but with all terms multiplied by a factor Q^p .

My entire series of papers deals with the optical problems strictly in the sense of A. H. Lorentz' Theory of Electrons, but differs from his treatment by the fact that no averaging over random positions of dipoles becomes necessary. This is the virtue of Sommerfeld's clear formulation of the problem. Shortly after my work, and independent of it, Professor C. W. Oseen in Oslo found the same screening action of a boundary that is so essential in my treatment. His work is based on Maxwell's equations for a homogeneous medium. He assumes a light wave to fall on a finite body bounded by a surface S , on which the usual boundary conditions hold. By transforming the integral expressing the polarization created in the body, he shows that this consists of the two parts, one of which conforms to the wave equation $\Delta u + \kappa^2 u = 0$ with $\kappa = K_0 = nk_0$ and the other with $\kappa = k_0$, where k_0 = wave constant in free space, n = refractive index and K_0 = wave constant in the medium. The same holds for the total field, which is the field created by the volume polarization in the body plus the field incident from outside. The condition of self-consistency in the body can be fulfilled only if those parts of the field vanish that progress with velocity c . That is, the external field into which the body is placed, must cancel that

part of the field which originates in the volume polarization and progresses with velocity c . The latter part can be expressed as a surface integral of the polarization over the boundary surface S . Thus it is by the action of the surface that the body is shielded from the impinging optical field.

This is in general terms the same result that I obtained for the special case of a plane boundary. The Oseen theory is well rendered elsewhere.*

Looking back on the mathematical technique used in Parts I and II, it may well be considered cumbersome. When the papers were written, the technique of Fourier transformation was still in its infancy. Reciprocal sets of axes had yet to be discovered by physicists. Their importance for the Fourier development of functions in a non-orthogonal cell was shown in my paper** on the reciprocal lattice, but even then the idea of considering the Fourier coefficients as the transform of a periodic function had either not been pronounced or not gained the attention of the physicists. Nowadays, the transformation of the sum of spherical dipole waves into a sum of plane waves is achieved easily by applying a few general properties of Fourier transformation. This method can be found in my paper,*** recording a lecture given in 1932.

My papers contain, for the first time I believe, a clear separation of the internal problem of finding the kind of optical field that can exist inside a dispersive medium, from the external problem of finding out what happens if an external wave falls on the surface of such a medium. It is only lately that I have become aware that the splitting of the entire problem into these two parts corresponds quite closely to the division into two steps of the mechanical problem of oscillations generated by an impact on a mechanical system. There, a preliminary study is made of the free oscillations of the system, and the knowledge so gained is subsequently applied to the behavior

* Born and Wolf, Principles of Optics, 2nd ed. (1964) pp. 100-104, Pergamon Press. The original paper by Oseen is Ann. d. Phys. (Lpz.), (1915) 48:pp. 1-56.

**Zeitsch. f. Kristallog (1921) 56:129-156.

***Annales de l'institut Henri Poincaré, 8:79-110.

of the system after an impact has set it in motion. The free oscillations are those requiring no outside forces and they can exist only at certain frequencies. The theory of dispersion (Part I) omits all external forces; to the determination of the proper frequency (or frequencies) of the mechanical system there corresponds the determination of the wave vector K (or, in the case of X-rays of the several wave vectors $K_b^{(i)}$) for which alone the system of field and dipoles is self-consistent. This knowledge is used to construct the state of the system filling a half-space when an impact is given to it by an external wave impinging on its surface. The development of the mechanical system in time is paralleled by the spatial development in the optical system. This analogy will be found worked out, for the case of X-rays, in a recent paper of mine.*

It is deplorable that most physicists interested in optics, as a sub-division of Physics, ignore or nearly ignore the fascinating branch of X-ray optics. In my papers, visible and X-ray optics are joined together in a comprehensive theory which, because of the mathematical strictness of its foundation, holds for the whole range of wavelengths, from the static case (λ infinite), through the visible ($\lambda \sim 5000 \text{ \AA}$), to X-rays ($\lambda \sim 1 \text{ \AA}$). The discussion, it is true, varies in these different regions, but the principles on which the theory is based remain the same. This wide range has attracted me to the theory through about 60 years in the course of which the X-ray optics gradually developed from a purely logical necessity to a reliable guide through a wealth of experimental observations which became feasible with the advent of the art of growing nearly faultless crystals.

* Acta Cryst. (1969) A25:103-108.

Appendix 8

Biography of P.P. Ewald

Paul P. Ewald was born 23 January 1888 in Berlin, Germany. His father, a promising young historian, died of appendicitis at the age of 37 shortly before he was born. His mother, being free to travel, took him on extended trips to England and to Paris. Thus by the age of five he had a rudimentary knowledge of both English and French.

He graduated from the Victoria Gymnasium of Potsdam obtaining his Abitur (school-leaving examination) in the fall of 1905. The two winter terms 1905/06 were spent at Gonville and Caius College, part of the University of Cambridge, England; 1906/07 at the University of Göttingen. From 1907 to 1912, he studied under Professor Arnold Sommerfeld at the University of Munich and received his Ph. D. in Physics in February 1912. His choice of a thesis subject influenced his entire scientific career. "Dispersion and Double Refraction of Electron Lattices (Crystals)" dealt with the propagation of light in crystals and was explored by him further in various ways in later papers.

The years of World War I were spent as an X-ray technician of the German army on the Russian front.

Returning to university life at the University of Munich in 1918, he became Privatdozent (lecturer) there, but soon moved to the Technische Hochschule, Stuttgart, in 1921 as Professor of Theoretical Physics, a

position he held until 1937. During this time he wrote a book: *Kristalle und Röntgenstrahlen* (Crystals and X-rays), (Springer 1923), and together with C. Hermann, *Strukturbericht 1913-1928* (Structure Report 1913-1928), (Akad, Verlagsges. 1931).

His first visit to the United States occurred in 1936 when he was invited to give a course of lectures on X-ray diffraction at the Summer School in Physics in Ann Arbor. Leaving Germany in 1937, he returned to Cambridge University, England, where from 1937 to 1939 he was a Research Fellow in the Crystallographic Laboratory. Then he moved to The Queen's University, Belfast, first as Lecturer, then as Professor of Mathematical Physics.

After ten years in Belfast, he moved to the United States and in 1949 became Professor of Physics and Head of the Physics Department at Polytechnic Institute of Brooklyn, Brooklyn, New York. In 1957 he relinquished the position as head of the physics department but remained as Professor of Physics until he retired in 1959.

Over his long and productive career he has written numerous papers of which he considers some to be fundamental. He is known for his part in the Ewald-Oseen extinction theorem, and Ewald's reciprocal lattice is a subject of study for students of optics and crystals. It is an important tool for all work on X-ray diffraction in crystals. Ewald's principal achievement is the "Dynamical Theory of X-ray Diffraction" which he dealt with in a number of papers using the same principles he had employed in his doctoral thesis and his two-part paper published in 1916 of which his thesis was Part I. These papers elucidate the principles which govern the transmission of electromagnetic radiation through a crystalline medium. The wavelength range which is considered includes both visible light and X-rays. Because of his work he has been the recipient of a number of honors. He was awarded an honorary doctorate from the Technische Hochschule, Stuttgart in 1954; University of Paris in 1958; Adelphi University, New York in 1966; and the University of Munich in 1967. He was made a Fellow of the Institute of Physics and Physical Society (England)

in 1945; the American Physical Society, 1952; the American Academy of Arts and Sciences, 1955; The Royal Society, 1958; Akademie Deutscher Naturforscher (Leopoldina) 1966.

He is corresponding member of the Akademie der Wissenschaften zu Göttingen (1937) and of the Bayerische Akademie der Wissenschaften in München, (1962). He was made an honorary member of Société française de minéralogie et cristallographie in 1955; Deutsche Mineralogische Gesellschaft 1958; Cambridge Philosophical Society, 1969.

Over the years, in addition to his numerous papers, he found time to be a co-editor of Zeitschr. f. Kristallographie from 1922 to 1937; and to be editor of Acta Crystallographica from its earliest beginnings in 1948 to 1959. In addition to the books previously mentioned, he contributed sections to: Landolt-Börnstein Tabellen (two editions); Müller-Pouillet Lehrbuch d. Physik (11th Edition 1926) (An English translation is in "The Physics of Solids and Fluids" (1st Ed. 1930, 2nd Ed. 1936) (Blackie)); Handbuch d. Physik (Springer 1927 and 1933); High Speed Aerodynamics Vol. 1 (Editor, F. D. Rossini, Princeton University Press 1955). He was the principal author as well as editor of the book "Fifty Years of X-ray Diffraction" published by the International Union of Crystallography in July 1962 to celebrate the Laue and Bragg discoveries.

He was Member of the executive Committee of the International Union of Crystallography from 1948 to 1966, and President of the same organization from 1960 to 1963.

The January 1968 issue of Acta Crystallographica was dedicated to him on the occasion of his 80th birthday. The issue contains one of his scientific articles and a three page article with the title "Personal Reminiscences", which contains many interesting biographic remembrances.

His retirement from university life in 1959 was not a retirement from other activities. True, he moved to quieter surroundings in New Milford, Connecticut, but as indicated by dates already given he has remained active. For a number of years he maintained a rather busy schedule of lectures in addition to his activities for the International Union of Crystallography.

B-1

Presently, (May 1970), he continues an active interest in science and carries on a scientific correspondence of some magnitude and occasionally travels to a meeting even if it is far away from Connecticut.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D		
(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified.)		
1. ORIGINATING ACTIVITY (Corporate author) Air Force Cambridge Research Laboratories (CA) L.G. Hanscom Field Bedford, Massachusetts 01730		20. REPORT SECURITY CLASSIFICATION Unclassified
		25. GROUP
3. REPORT TITLE ON THE FOUNDATIONS OF CRYSTAL OPTICS Part I - Dispersion Theory Part II - Theory of Reflection and Refraction		
4. DESCRIPTIVE NOTES (Type of report and previous dates) Scientific Interim		
5. AUTHOR (Last name, first initial, last name) P.P. Ewald		
6. REPORT DATE 26 August 1970	75. TOTAL NO. OF PAGES 110	75. NO. OF REFS 6
10. CONTRACT OR GRANT NO.		20. ORIGINATOR'S REPORT NUMBER(S)
a. PROJECT, TASK, WORK UNIT NOS.	N/A	AFCL-70 3520
c. DOD ELEMENT	N/A	95. OTHER REPORT NUMBER(s) (Any other numbers that may be assigned this report) Translation No. 54
d. DOD SUBELEMENT	N/A	
10. DISTRIBUTION STATEMENT 1 - This document has been approved for public release and sale; its distribution is unlimited.		
11. SUPPLEMENTARY NOTES Translation: From Annalen der Physik, 49, 1915, No. 1 and No. 2. Postscript 1970 by P.P. Ewald KEYWORDS: Dispersion theory, Reflection, Refraction		12. SPONSORING MILITARY ACTIVITY Air Force Cambridge Research Laboratories (CA) L.G. Hanscom Field Bedford, Massachusetts 01730
13. ABSTRACT This two-part paper investigates the propagation of light in the visible region through a crystalline medium. In Part I, which is a slightly shortened version of the author's doctoral thesis of 1912, the objective is to determine if the anisotropic arrangement of ordinary (isotropic) dipoles at the axial points of an orthorhombic lattice would account for the existence of double refraction. The value of the computed birefringence is found to be comparable to the observed value. Additionally some features of the traditional "theory of dispersion" are disclosed and clarified. In the older theories the "incident optical wave" plays a role but it is shown that in a medium which extends to infinity in all directions no such wave should be assumed. The refractivity is shown to arise as an internal property of the medium. In Part II, a crystalline medium is considered as filling a half-space and having a plane boundary at $x = 0$. A plane optical wave is incident on this medium. Because of the linearity of the equations it has to be superimposed on the field originating in the crystal. It is shown that this incident optical wave is actually prevented from entering the crystal because of the modification produced in the field of the crystal by the introduction of a boundary. Boundary waves are found to exist on both sides of the boundary. The higher their order, the more rapidly they attenuate as a function of distance away from the boundary but the zero order waves are ordinary undisturbed plane waves of vacuum velocity c . The fields outside and inside the medium are connected by the Fresnel formulas. These follow from the condition that the optical field inside the medium and the oscillations generated there form a self-consistent system. Material Added 1970. The conclusion that the incident optical field cannot penetrate the crystal boundary, together with a similar conclusion in a paper by Oseen, is the basis of the Fernald-Oseen Extinction Theorem. The same methods and field transformations developed in this two part paper were applied in two later papers which were published as Parts III and IV. These extend the treatment to the passage of X-rays through crystals. "Postscript 1970" published here indicates the relationship of these four parts and further developments of the subject as well as referencing the papers noted above.		

DD FORM 1073

Unclassified

Security Classification

Unclassified

Security Classification

14.	KEY WORDS	LINK A		LINK B		LINK C	
		ROLE	WT	ROLE	WT	ROLE	WT
	Dispersion theory Reflection Refraction						

Unclassified

Security Classification