Reprinted from *J. Opt. Soc. Am.*, **31**, 488–493 (July 1941)

A New Calculus for the Treatment of Optical Systems

I. Description and Discussion of the Calculus

R. CLARK JONES*

Research Laboratory, Polaroid Corporation, Cambridge, Massachusetts, and Research Laboratory of Physics, Harvard University, Cambridge, Massachusetts

(Received April 23, 1941)

The effect of a plate of anisotropic material, such as a crystal, on a collimated beam of polarized light may always be represented mathematically as a linear transformation of the components of the electric vector of the light. The effect of a retardation plate, of an anisotropic absorber (plate of tourmaline; Polaroid sheeting), or of a crystal or solution possessing optical activity, may therefore be represented as a matrix which operates on the electric vector of the incident light. Since a plane wave of light is characterized by the phases and amplitudes of the two transverse components of the electric vector, the matrices involved are two-by-two matrices, with matrix elements which are in general complex. A general theory of optical systems containing plates of the type mentioned is developed from this point of view.

INTRODUCTION

THE type of optical system with which this paper deals is not the more familiar type involving lenses, prisms, etc., but is rather the type which is composed of retardation plates, partial polarizers, and plates possessing the ability to rotate the plane of polarization. We shall therefore be concerned not with the directions of rays of light, but with the state of polarization and the intensity of the light as it passes through the optical system.

An example of a partial polarizer is a plate of tourmaline, or a sheet of light-polarizing film such as those sold under the trade-name "Polaroid." [1] For light incident perpendicularly on the plate, a partial polarizer may be characterized by two mutually perpendicular axes, with each of which is associated one of the two principal absorption coefficients. The retardation plate, or wave plate, is too well known to require description. In this paper we shall refer to a plate having the ability to rotate the plane of polarization as a *rotator*. A rotator is exemplified by a plate of quartz viewed along the optic axis, or by solutions of optically active molecules. We shall refer to the individual plates which make up the complete optical system as the elements of the system, or as the optical elements.

In the present paper it will be assumed that the light is always incident normally on the elements of the optical system.

The present paper is devoted to a description of the new calculus. When light passes through any of the three fundamental types of optical elements, the state of polarization and sometimes the intensity of the light will be changed. We shall find it possible to represent the effect of any optical element on the light as a linear operator acting upon the electric vector of the light wave. The operator is expressed in the convenient form of a two-by-two matrix, whose four matrix elements are in the general complex. From the associative property of matrices, it will then follow that the complete optical system may also be represented by a two-by-two complex matrix.

In Part II, the calculus is used to prove three theorems relating to optical systems of the type considered. The two theorems relating to systems containing partial polarizers are new, so far as the writer knows, and should be of some practical importance because of the recent development of a readily available and high quality polarizer in the form of Polaroid sheeting.

In Part III, the theory is applied to a rigorous treatment of the Sohncke theory of optical activity.

* National Scholar, Harvard University.

[1] Registered trade-mark of the Polaroid Corporation. For a description of Polaroid sheeting, see Martin Grabau, J. Opt. Soc. Am. **27**, 420 (1937).

Conventions and Notation

Consider a right-handed rectangular coordinate system, x, y and z. The elements of the optical system are considered to be arranged along the z axis, the z axis being perpendicular to the plane of the plates. The optical system will always be described as viewed from a point which is further out along the positive z axis than any of the elements; with this convention, the x and y axes have the relative position of the usual xy plane: the positive x axis may be superposed upon the positive y axis by rotating the x axis 90° counterclockwise.

Furthermore, let the two principal axes of the ith retardation plate or partial polarizer be indicated by x_i' and y_i'. We assume that the positive x_i' and y_i' axes have the same relative orientation as the positive x and y axes. We may now define the orientation of the ith element by stating the angle ω_i measured counterclockwise from the positive x axis to the positive x_i axis.

The light will be represented as plane waves progressing in either direction along the z axis. The state of polarization may be completely defined by stating the amplitudes and phases of the x and y components of the electric vector of the light wave. At any fixed point along the z axis, the components may be written in the usual complex form

$$\begin{aligned} E_x &= A_x \exp\left[i(\epsilon_x + 2\pi\nu t)\right], \\ E_y &= A_y \exp\left[i(\epsilon_y + 2\pi\nu t)\right], \end{aligned} \tag{1}$$

where A_x and A_y, ϵ_x and ϵ_y are real. If $\epsilon_x - \epsilon_y$ is an integral multiple of π, the light is plane polarized; otherwise, elliptically polarized.

Now let us consider the change in the character of a light wave as it passes through a retardation plate or a partial polarizer. As a matter of fact, in order to avoid giving a separate treatment of the two cases, we shall suppose that both the indices of refraction and the absorption coefficients are different along the two principal axes (x' and y') of the plate. If we know the x' and y' components ($E_{x'0}$ and $E_{y'0}$) of the electric vector as the light enters the plate, then the corresponding components of the light as it emerges from the other side of the plate are

$$\begin{aligned} E_{x'1} &= E_{x'0} \exp\left[-i(2\pi d/\lambda)(n_{x'} - ik_{x'})\right] \\ &= N_{x'} E_{x'0}, \\ E_{y'1} &= E_{y'0} \exp\left[-i(2\pi d/\lambda)(n_{y'} - ik_{y'})\right] \\ &= N_{y'} E_{y'0}. \end{aligned} \tag{2}$$

In these expressions, d is the thickness of the plate, λ is the wave-length of the light in vacuum, the n's are the principal indices of refraction, and the k's are the principal extinction coefficients. The extinction coefficient is a particular type of amplitude absorption coefficient. The N's are merely abbreviations for the exponential expressions. In the case of a retardation plate, the k's are the same and the n's different, whereas in the case of a partial polarizer, the n's are the same and the k's different. In practice, of course, it is difficult to secure a partial polarizer which is not also birefringent.

In general, however, we are interested in knowing the change in the x and y components of the light wave, rather than the change in the x' and y' components. If ω is the angle measured counterclockwise from the positive x axis to the positive x' axis, we have the relations

$$\begin{aligned} E_{x'} &= E_x \cos\omega + E_y \sin\omega, \\ E_{y'} &= -E_x \sin\omega + E_y \cos\omega. \end{aligned} \tag{3}$$

The elimination of the x' and y' components in (2) and (3) yields the result

$$\begin{aligned} E_{x1} &= (N_{x'} \cos^2\omega + N_{y'} \sin^2\omega) E_{x0} \\ &\quad + (N_{x'} - N_{y'}) \sin\omega \cos\omega E_{y0}, \\ E_{y1} &= (N_{x'} - N_{y'}) \sin\omega \cos\omega E_{x0} \\ &\quad + (N_{x'} \sin^2\omega + N_{y'} \cos^2\omega) E_{y0}. \end{aligned} \tag{4}$$

The equations (4) give the important relation between the x and y components of the entering and emergent light.

The Matrix Notation

The relations (4) may be put in a convenient and simple form by the use of a matrix notation.[2]

[2] In Parts I and II we shall need only the algebraic properties of matrices, as they are presented on pp. 348–352 of E. C. Kemble, *Fundamental Principles of Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1937). In Part III, we shall find it necessary to use also the transformation properties of matrices; see the reference just given, or V. Rojansky, *Introductory Quantum Mechanics* (Prentice-Hall, New York, 1938), pp. 285–340.

Let **M** be the two-by-two matrix*

$$\mathbf{M} \equiv \begin{pmatrix} m_1 & m_4 \\ m_3 & m_2 \end{pmatrix}, \tag{5}$$

where

$$\begin{aligned} m_1 &= N_{x'} \cos^2 \omega + N_{y'} \sin^2 \omega, \\ m_2 &= N_{x'} \sin^2 \omega + N_{y'} \cos^2 \omega, \\ m_3 &= m_4 = (N_{x'} - N_{y'}) \sin \omega \cos \omega. \end{aligned} \tag{6}$$

Furthermore, let ε_0 and ε_1 be the one-column vectors

$$\varepsilon_0 \equiv \begin{pmatrix} E_{x0} \\ E_{y0} \end{pmatrix}; \quad \varepsilon_1 \equiv \begin{pmatrix} E_{x1} \\ E_{y1} \end{pmatrix}. \tag{7}$$

The relations (4) are now equivalent to the vector equation

$$\varepsilon_1 = \mathbf{M}\varepsilon_0. \tag{8}$$

The matrix **M** may itself be written in a simple form. Let $\mathbf{S}(\omega)$ be the rotation matrix

$$\mathbf{S}(\omega) \equiv \begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix} \tag{9}$$

and let

$$\mathbf{N} \equiv \begin{pmatrix} N_{x'} & 0 \\ 0 & N_{y'} \end{pmatrix}. \tag{10}$$

We then have

$$\mathbf{M} = \mathbf{S}(\omega)\mathbf{N}\mathbf{S}(-\omega) \tag{11}$$

or

$$\varepsilon_1 = \mathbf{S}(\omega)\mathbf{N}\mathbf{S}(-\omega)\varepsilon_0. \tag{12}$$

We have thus been able to represent the effect of a retardation plate or a partial polarizer as a matrix operator which operates upon the vector describing the intensity and polarization of the entering light. The matrix operator has been factored into a product of two types of matrices, the first of which, **N**, describes the optical element in a way independent of its orientation, and the second of which, **S**, describes the orientation.

Multi-Element Systems

Suppose now that we have a series of n optical elements, represented by the n matrices $\mathbf{N}_1$, $\mathbf{N}_2$, $\cdots$, $\mathbf{N}_n$, with the respective orientations ω_1, ω_2, $\cdots$, ω_n, and suppose that the light passes through the elements in the order in which they are numbered. We then have

$$\begin{aligned} \varepsilon_1 &= \mathbf{S}(\omega_1)\mathbf{N}_1\mathbf{S}(-\omega_1)\varepsilon_0 = \mathbf{M}_1\varepsilon_0, \\ \varepsilon_0 &= \mathbf{M}_2\varepsilon_1, \\ &\cdots \\ \varepsilon_n &= \mathbf{M}_n\varepsilon_{n-1}, \end{aligned} \tag{13}$$

where ε_i is the light vector as it emerges from the ith element. By substituting each of the Eqs. (13) in the one following it, we have at once the relation between ε_n and ε_0

$$\begin{aligned} \varepsilon_n &= \mathbf{M}_n\mathbf{M}_{n-1}\cdots\mathbf{M}_2\mathbf{M}_1\varepsilon_0 \\ &= \mathbf{M}^{(n)}\varepsilon_0, \end{aligned} \tag{14}$$

where $\mathbf{M}^{(n)}$ is the two-by-two matrix

$$\mathbf{M}^{(n)} \equiv \mathbf{M}_n\mathbf{M}_{n-1}\cdots\mathbf{M}_2\mathbf{M}_1. \tag{15}$$

According to (6), each of the $\mathbf{M}_i$'s is a symmetrical matrix. In general, however, the product of two or more of the **M** matrices will be neither symmetrical nor anti-symmetrical. This criterion permits us to distinguish immediately between the matrices corresponding to a single element, and those corresponding to the superposition of two or more elements.

It is fruitful to examine the structure of $\mathbf{M}^{(n)}$ more closely. We have

$$\mathbf{M}^{(n)} = [\mathbf{S}(\omega_n)\mathbf{N}_n\mathbf{S}(-\omega_n)][\mathbf{S}(\omega_{n-1})\mathbf{N}_{n-1}\mathbf{S}(-\omega_{n-1})]\cdots[\mathbf{S}(\omega_2)\mathbf{N}_2\mathbf{S}(-\omega_2)][\mathbf{S}(\omega_1)\mathbf{N}_1\mathbf{S}(-\omega_1)]. \tag{16}$$

By the use of the easily proved relation:

$$\mathbf{S}(\omega_1+\omega_2) \equiv \mathbf{S}(\omega_1)\mathbf{S}(\omega_2), \tag{17}$$

the expression for $\mathbf{M}^{(n)}$ may be written in either of the following forms:

$$\mathbf{M}^{(n)} = \mathbf{S}(\omega_n)[\mathbf{N}_n\mathbf{S}(\omega_{n,\,n-1})\cdots\mathbf{N}_2\mathbf{S}(\omega_{2,\,1})\mathbf{N}_1\mathbf{S}(\omega_{1,\,n})]\mathbf{S}(-\omega_n), \tag{18}$$

$$= \mathbf{S}(\omega_1)[\mathbf{S}(\omega_{1,\,n})\mathbf{N}_n\mathbf{S}(\omega_{n,\,n-1})\cdots\mathbf{S}(\omega_{2,\,1})\mathbf{N}_1]\mathbf{S}(-\omega_1), \tag{19}$$

where

$$\omega_{i,\,j} \equiv \omega_j - \omega_i. \tag{20}$$

* We depart from the usual double-subscript notation for matrix elements solely for reasons of typographical convenience.

The matrices represented by the square brackets in Eqs. (18) and (19) depend only on the relative orientation of the elements of the optical system, and are independent of the orientation of the optical system as a whole. Thus, just as in the case of a single element, we have factored the operator representing the optical system into the product of two types of matrices, the first of which depends only on the nature of the optical elements and on their *relative* orientation, and the second of which describes only the orientation of the optical system as a whole with respect to the x, y axes.

So far we have excluded any consideration of rotators, which are plates possessing the ability to rotate the plane of polarization about the z axis. The formal change caused by the introduction of rotators into the system is very small, however. Suppose that a given rotator is able to rotate the plane of polarization through the angle $\bar{\omega}$ in the counterclockwise direction. The matrix operator corresponding to the rotator is then simply $\mathbf{S}(\bar{\omega})$. We shall not stop to prove this statement,* since its proof is elementary. If the rotator just considered is introduced between the ith and the $(i+1)$th element of the optical system considered in the last paragraph, the resulting optical system may still be represented in the form of Eqs. (18) and (19) if we redefine $\omega_{i+1,\, i}$ as

$$\omega_{i+1,\, i} \equiv \omega_i - \omega_{i+1} + \bar{\omega}. \tag{21}$$

The generalization to the case in which several rotators are placed in the optical system is obvious. If several rotators are placed in juxtaposition, they are equivalent to a single rotator, according to (17).

There is an important physical distinction between the behavior of retardation plates and partial polarizers on one hand, and the behavior of rotators on the other. With the coordinate system which we are using, it makes no difference in what direction the light passes through a retardation plate or partial polarizer. With the type of rotator represented by a crystal or solution possessing optical activity, however, the rotation changes sign if the light passes through the plate in the reverse direction. In determining the operator corresponding to a rotator, therefore, it is necessary to know in which direction the light passes through the plate.

* More precisely, the matrix of a rotator is $e^{i\phi}\mathbf{S}(\bar{\omega})$, where the phase factor takes into account the finite optical thickness of the plate. In this paper, however, rotators will be considered only in combination with retardation plates and partial polarizers, so that the phase factor may always be considered to be associated with the other elements of the optical system. For simplicity, we omit the phase factor.

Reversibility of the Optical System

It is, however, not necessary to make a recalculation if the light passes through the optical system in the opposite direction. Suppose that we have computed the matrix $\mathbf{M}^{(n)}$ for the case that the light passes through the system in the order in which the plates are numbered, and then suppose that the light is reversed so that it passes through the system in the opposite direction, passing first through the nth plate, and last through the first plate. In terms of the one-row vectors,

$$\bar{\varepsilon} \equiv (E_x\ E_y), \tag{22}$$

the relation between the vector $\bar{\varepsilon}_{\text{initial}}$ of the light entering the nth plate and the vector $\bar{\varepsilon}_{\text{final}}$ of the light emerging from the first plate is

$$\bar{\varepsilon}_{\text{final}} = \bar{\varepsilon}_{\text{initial}} \mathbf{M}^{(n)}, \tag{23}$$

where $\mathbf{M}^{(n)}$ is exactly the same matrix as that previously defined. This statement may be proved easily, first by showing that a relation of the form (23) holds for each element of the system, and then by eliminating the intermediate $\bar{\varepsilon}$ vectors.

The statement made in the last paragraph holds only if all of the rotators in the system are of the type represented by ordinary optical activity, and not of the type represented by the Faraday effect, in which case the direction of the rotation depends on the direction of the magnetic field, and not on the direction in which the light passes through the material. If any of the rotators are of the type represented by the Faraday effect, then the $\mathbf{M}^{(n)}$ which appears in (23) is not the same as the $\mathbf{M}^{(n)}$ which appears in (14), the difference being that the $\bar{\omega}$'s appearing in the $\mathbf{S}$ matrices which represent the Faraday rotators must have opposite signs in the two $\mathbf{M}^{(n)}$'s.

Discussion of the Matrices

When the optical system does not contain any partial polarizers but consists entirely of rotators and retardation plates, *all of the matrices are unitary.*‡ This fact may be proved by either of two methods.

First proof

The **N** matrices are all unitary, since their characteristic values are all of absolute value unity; the latter is a necessary and sufficient condition that a matrix be unitary. Furthermore, the **S** matrices are easily shown to be unitary, since the conjugate transpose of $\mathbf{S}(\omega)$ is $\mathbf{S}(-\omega)$, so that $\mathbf{S}^\dagger\mathbf{S}=1$. The general matrix $\mathbf{M}^{(n)}$ is the product of factors which are all unitary, and therefore $\mathbf{M}^{(n)}$ is unitary.

Second proof

This proof is not purely mathematical, but is based on the fact that an optical system composed only of rotators and retardation plates does not change the intensity of the light passing through it. The intensity of the light is proportional to the sum of the squares of the x and y components of the electric vector:

$$I \propto A_x^2+A_y^2=E^*_xE_x+E^*_yE_y=\tilde{\varepsilon}^*\varepsilon=\tilde{\varepsilon}\varepsilon^* \quad (24)$$

(cf. Eq. (1) for part of the notation). The intensity of the light is therefore proportional to the square of the length of the vector ε. Now a unitary matrix does not change the length of a vector, and conversely, if a matrix does not change the length of an arbitrary vector, then the matrix is unitary. Thus, since the optical system does not change the intensity of the light, the corresponding matrices must be unitary.

Conversely, if the first proof of the unitary character of the matrices is accepted, then the second may be regarded as a proof that an optical system composed only of retardation plates and rotators does not change the intensity of the light passing through it.

In nearly every case, we are not interested in the total phase change corresponding to the two axes of a retardation plate, but only in the phase difference of the two principal axes; we often refer to a retardation plate as a q wave plate, implying that the optical path difference for plane polarized light parallel to each of the two axes is q wave-lengths.

Since the two principal extinction coefficients of a retardation plate are both zero, we find by comparison of (2) with (10) that the matrix of a retardation plate whose axes are parallel to the x and y axes may always be written in the form $e^{i\phi}\mathbf{G}$, where

$$\mathbf{G}\equiv\begin{pmatrix} e^{i\gamma} & 0 \\ 0 & e^{-i\gamma} \end{pmatrix}. \quad (25)$$

If we are interested only in the phase difference, the phase factor $e^{i\phi}$ may be omitted, and we may consider **G** alone to be the diagonal representation of a wave plate.

If we choose to write the diagonal matrices of all the wave plates in the optical system in the form (25), then, since the determinant of an **S** matrix is unity, all of the matrices which occur as factors on the right-hand side of Eq. (16) will have determinants equal to unity, provided that the system contains only retardation plates and rotators. Then, since the determinant of a product of matrices is equal to the product of the determinants of the matrices, the determinant of $\mathbf{M}^{(n)}$ will have the value unity.

Our treatment of partial polarizers is similar. The diagonal form of the matrix corresponding to a partial polarizer may always be written in the form $e^{i\phi}\mathbf{P}$, where

$$\mathbf{P}\equiv\begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix} \quad \begin{matrix} 0\leqslant p_1\leqslant 1 \\ 0\leqslant p_2\leqslant 1. \end{matrix} \quad (26)$$

Again, if we are not interested in the absolute phase of the two components of the light, but only in the relative phase, we may consider **P** alone to be the diagonal representation of a partial polarizer. If the plate is a perfect polarizer, that is to say, if the plate absorbs completely light parallel to one of its axes, then one of the two p's will be zero; in this case the determinant of **P** is zero. More generally, the determinant of **P** is real, with its value lying in the range zero to unity.

We thus find that if we use **G** to represent a retardation plate, and **P** to represent a partial polarizer, then the determinant of $\mathbf{M}^{(n)}$ for an optical system containing all three types of

‡ Let $\mathbf{M}^\dagger$ be the matrix obtained from **M** by first exchanging m_3 and m_4 and then replacing each of the matrix elements by its complex conjugate. The matrix **M** is unitary if $\mathbf{M}^\dagger\mathbf{M}=\mathbf{M}\mathbf{M}^\dagger=1$. The matrix $\mathbf{M}^\dagger$ is said to be the Hermitian adjoint, or the conjugate transpose of **M**.

elements will be real and non-negative, with its value in the range zero to unity.

It will be shown in Part II that the most general unitary matrix may be represented by an optical system consisting of only two elements, one a rotator and the other a retardation plate. Similarly, it will be shown that the general two-by-two complex matrix may be represented by an optical system containing not more than four elements, the only restriction being that the matrix may not be one which would increase the intensity of the light.

Elliptical Polarization

For convenience in reference, we state here the relations giving the shape and orientation of the ellipse representing the state of polarization of the light. Let the amplitude of the x and y components be A_x and A_y, and let $\delta=\epsilon_y-\epsilon_x$, as in Eq. (1). Further, let $\tan\theta$ equal the ratio of axes of the ellipse, and let ψ be the orientation of the major axis of the ellipse, measured counter-clockwise from the x axis. The following relations determine θ and ψ:[3]

$$\begin{aligned}\tan 2\psi&=\tan 2\alpha\cos\delta,\\ \cos 2\theta&=\sin 2\alpha|\sin\delta|,\end{aligned}\tag{27}$$

where

$$\tan\alpha=A_y/A_x,\quad 0\leqslant\alpha\leqslant\tfrac{1}{2}\pi.\tag{28}$$

When $\sin\delta$ is positive, the ellipse is described in the clockwise direction, and *vice versa*.

[3] Max Born, *Optik* (Springer, Berlin, 1933), p. 23.