THE PHYSICS OF POLARIZATION

By EGIDIO LANDI DEGL'INNOCENTI

Dipartimento di Astronomia e Scienza dello Spazio, Università di Firenze, Largo E. Fermi 5, 50125 Firenze, Italy

This course is intended to give a description of the basic physical concepts which underlie the study and the interpretation of polarization phenomena. Apart from a brief historical introduction (Sect. 1), the course is organized in three parts. A first part (Sects. 2-6) covers the most relevant facts about the polarization phenomena that are typically encountered in laboratory applications and in everyday life. In Sect. 2, the modern description of polarization in terms of the Stokes parameters is recalled, whereas Sect. 3 is devoted to introduce the basic tools of laboratory polarimetry, such as the Jones calculus and the Mueller matrices. The polarization phenomena which are met in the reflection and refraction of a beam of radiation at the separation surface between two dielectrics, or between a dielectric and a metal, are recalled in Sect. 4. Finally, Sect. 5 gives an introduction to the phenomena of dichroism and of anomalous dispersion and Sect. 6 summarizes the polarization phenomena that are commonly encountered in everyday life. The second part of this course (Sects. 7-14) deals with the description, within the formalism of classical physics, of the spectro-polarimetric properties of the radiation emitted by accelerated charges. Such properties are derived by taking as starting point the Liénard and Wiechert equations that are recalled and discussed in Sect. 7 both in the general case and in the non-relativistic approximation. The results are developed to find the percentage polarization, the radiation diagram, the cross-section and the spectral characteristics of the radiation emitted in different phenomena particularly relevant from the astrophysical point of view. The emission of a linear antenna is derived in Sect. 8. The other Sections are devoted to Thomson scattering (Sect. 9), Rayleigh scattering (Sect. 10), Mie scattering (Sect. 11), bremsstrahlung radiation (Sect. 12), cyclotron radiation (sect. 13), and synchrotron radiation (Sect. 14). Finally, the third part (Sects. 15-19) is devoted to give a sketch of the theory of the generation and transfer of polarized radiation in spectral lines. After a general introduction to the argument (Sect. 15), the concepts of density-matrix and of atomic polarization are illustrated in Sect. 16. In Sect. 17, a parallelism is established, within the framework of the theory of stellar atmospheres, between the usual formalism, which neglects polarization phenomena, and the more involved formalism needed for the interpretation of spectro-polarimetric observations. Some consequences of the radiative transfer equations for polarized radiation, pointing to the importance of dichroism phenomena in establishing the amplification condition via stimulated emission, are discussed in Sect. 18. The last section (Sect. 19) is devoted to introduce the problem of finding a self-consistent solution of the radiative transfer equations for polarized radiation and of the statistical equilibrium equations for the density matrix (non-LTE of the 2nd kind).

1. Introduction

Polarization is an important physical property of electromagnetic waves which is connected with the transversality character, with respect to the direction of propagation, of the electric and magnetic field vectors. Under this respect, the phenomenon of polarization is not restricted to electromagnetic waves, but could in principle be defined for any wave having a transverse character, such as, for instance, transverse elastic waves propagating in a solid, transverse seismic waves, waves in a guitar string, and so on. On the contrary, polarization phenomena are obviously inexistent for longitudinal waves, such as the usual acoustic waves propagating in a gas or in a liquid.

From an historical perspective (see Swyndell, 1975, for a more exhaustive treatment of the argument), the study of the polarization characteristics of electromagnetic waves started as early as the 17th century with an interesting treatise by the Dutch physicist Erasmus Bartholinus entitled "*Experimenta crystalli islandici disdiaclastici, quibus mira et insolita refractio detegitur*" ("*Experiments on double-refracting Icelandic crystals, showing amazing and unusual refraction*", 1670 –detailed references to this work, as well as to the other papers or books appeared earlier than 1850, can be found in Swyndell, 1975). In this work, one can find the earliest account of a phenomenon, double refraction in a crystal, which is intimately connected with the polarization characteristics of light. We now know that the two rays resulting from the refraction inside a crystal such as an Iceland spar have different polarization characteristics, a fact that was however ignored by Bartholinus.

Reflecting on Bartholinus' experiments, first Christian Huyghens in his "*Treatise on Light*" (1690), and later Isaac Newton in his "*Optiks*" (1730), though working in the framework of two competing theories of light, arrived to the conclusion that light should have some "transversality" property, a property, however, that was not yet called "polarization". Newton, for instance, refers to the phenomenon of polarization by saying that a ray of light has "sides".

After many years from Huyghens and Newton, the French physicist Etienne Louis Malus introduces in the scientific literature the word "*Polarization*" and brings several significant contributions to the establishment of the concept of polarization in modern terms. In his paper "*Sur une proprieté de la lumière réfléchie*" (1809) Malus proves that polarization is an intrinsic property of light (and not a property "induced" in the light by crossing an Iceland spar), he demonstrates that polarization can be easily produced through the phenomena of reflection and refraction, and he also proves the famous $\cos^2\theta$ law (giving the fraction of the intensity transmitted by two polarizers crossed at an angle θ), nowadays known as Malus law. This work opens the way to the achievements of another physicist, probably the most renowned optician of all times, Augustin Fresnel, who definitely proves the transversality of light despite the widespread belief of the times according to which, the ether being a fluid, the light should be composed of longitudinal waves. Around 1830, in his paper "*Mémoires sur la réflexion de la lumière polarisée*", Fresnel proves his famous laws concerning the relationships among the polarization properties of the incident beam and the same properties of the beams reflected and refracted at the surface of a dielectric. Despite the fact that the electromagnetic nature of light was not yet known, Fresnel's laws are correct and are still in use today.

The story of polarization continues in the 18th century with several significant contributions by François Arago and Jean-Baptiste Biot (who discover the phenomenon of *Optical Activity* in crystals and in solutions, respectively), David Brewster (nowadays known for the "Brewster angle"), William Nicol (who builds the first polarizer, the so-called Nicol prism), and Michael Faraday (who discovers an effect today known as the "Faraday effect"). However, it is only with the fundamental work of George Stokes, "*On the Composition and Resolution of Streams of Polarized Light from Different Sorces*" (1852), that the description of polarized radiation becomes fully consistent. This is achieved by giving an operational definition of four quantities, the so-called Stokes parameters, and by introducing a statistical description of the polarization property of radiation, as we will see in the next Section.

At the middle of the 18th century, the phenomenon of polarization is thus fairly well understood but it is necessary to wait almost 60 years before assisting to the first application of polarimetry to astronomy. In 1908, George Ellery Hale, has the brilliant idea of observing the solar spectrum with the help of some polarizing devices (Hale, 1908). By means of a Fresnel rhomb (acting as a quarter-wave plate) and a Nicol prism (acting as a polarizer), Hale succeeds in observing the spectrum of a sunspot in two opposite directions of circular polarization and, from the observed shift of spectral lines, induces for the first time the existence of magnetic fields in an astronomical object. Since its birth, astronomical polarimetry has evolved through the years and has given a relevant contribution to our present understanding of the physical Universe. Among the various astronomical discoveries that have relied on the use of polarimetric techniques it is just enough to quote here the discovery of the first magnetic star (Babcock, 1947) and the discovery of the existence of magnetic white dwarfs (Kemp *et al.*, 1970).

Notwithstanding these remarkable successes, polarimetry has remained for a long time a secondary discipline in astronomy. However, mostly in the last ten years, the situation has rapidly evolved and we are now undoubtedly assisting to a revival of this discipline that seems capable of capturing the scientific interests of a large community of persons and a non negligible fraction of the funds allocated to astronomical research (the organization of the present Winter School is a clear example of this trend). Probably, this is far from being an accidental event. Now that all the possible "windows" of the electromagnetic spectrum have been opened (from γ-rays to radio-waves), the possibility of new discoveries –including the serendipitous ones– relies on the development of new technologies aimed to increase the accuracy of older instrumentation (better angular, temporal, or spectral resolution, better photometric accuracy, and so on). Polarimetry perfectly fits into this trend also because, for almost a century, it has generally trailed behind the other disciplines as a possible target of novel technologies.

Apart from these historical notes, I feel necessary to spend some more introductory words about polarimetry in the astronomical context. The first thing to be remarked is that polarization is an invaluable source of information about *the geometry* of the astronomical object observed, or about any physical agent (like for instance a magnetic field) that is capable of altering, to some extent, the geometrical scenario of the same object. In polarimetry, more than in any other discipline of astronomy, the words of Galileo about geometry and the physical world still stand, after almost four centuries, as a must: "*Egli (l'Universo) è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali mezzi è impossibilie a intendere umanamante parola...*" ("*The Universe is written in mathematical language, and its characters are triangles, circles, and other geometrical figures, without which it is humanly impossible to understand a single word...*").

Only a perfectly symmetric object, devoided of any physical agent capable of introducing the minimum dissimmetry in its geometrical scenario, is capable of emitting a completely unpolarized beam of radiation. An ideal black-body could provide an example of such an object, but, as we all know, ideal objects do not exist in real life and we have then to expect that some polarization signal, even if exceedingly small, may always be present in no matter which astronomical object.

The real challenge for the future of astronomical polarimetry is to increase the sensitivity of the present polarimeters operating in the different regions of the electromagnetic spectrum. Quite recently, solar physicists have succeeded in lowering the sensitivity of their polarimeters, operating in the visible range of the electromagnetic spectrum, below the limit of 10^{-4}, thus discovering a wealth of new and unexpected phenomena that are taking place in the higher layers of the solar atmosphere and that are stimulating novel theoretical approaches for their interpretation. It is my impression that, quite similarly, new exciting discoveries may be obtained for any spectral domain and any discipline of astronomy once the major effort of building a new-technology polarimeter has reached the ultimate goal of lowering the sensitivity of presently available instruments.

2. Description of Polarized Radiation

Consider an electromagnetic, monochromatic plane wave of angular frequency ω that is propagating in vacuum along a direction that we assume as the z-axis of a right-handed reference system. In a given point of space, the electric and magnetic field vectors of the wave oscillate in the x-y plane according to equations of the form

$$E_x(t) = E_1 \cos(\omega t - \phi_1) \ , \qquad E_y(t) = E_2 \cos(\omega t - \phi_2) \quad ,$$

where E_1, E_2, ϕ_1, and ϕ_2 are constants. The same oscillation can also be described in terms of complex quantities by writing

$$E_x(t) = \mathrm{Re}(\mathcal{E}_1 \mathrm{e}^{-\mathrm{i}\omega t}) \ , \qquad E_y(t) = \mathrm{Re}(\mathcal{E}_2 \mathrm{e}^{-\mathrm{i}\omega t}) \quad ,$$

where $\mathcal{E}_1$ and $\mathcal{E}_2$ are given by

$$\mathcal{E}_1 = E_1 \mathrm{e}^{\mathrm{i}\phi_1} \ , \qquad \mathcal{E}_2 = E_2 \mathrm{e}^{\mathrm{i}\phi_2} \quad .$$

As is well known, the composition of two orthogonal oscillations of the same frequency gives rise to an ellipse. The tip of the electric field vector thus describes an ellipse at the angular frequency ω, and, when trying to recover the geometrical parameters of the ellipse from the quantities previously introduced, one finds that the following four combinations,

$$P_I = E_1^2 + E_2^2 = \mathcal{E}_1^* \mathcal{E}_1 + \mathcal{E}_2^* \mathcal{E}_2 \ , \quad P_Q = E_1^2 - E_2^2 = \mathcal{E}_1^* \mathcal{E}_1 - \mathcal{E}_2^* \mathcal{E}_2 \ ,$$

$$P_U = 2E_1 E_2 \cos(\phi_1 - \phi_2) = \mathcal{E}_1^* \mathcal{E}_2 + \mathcal{E}_2^* \mathcal{E}_1 \ , \quad P_V = 2E_1 E_2 \sin(\phi_1 - \phi_2) = \mathrm{i}\,(\mathcal{E}_1^* \mathcal{E}_2 - \mathcal{E}_2^* \mathcal{E}_1) \ ,$$

come naturally into play. The ratio between the minor and major axes of the ellipse, for instance, is given by

$$\frac{b}{a} = \frac{\left| \sqrt{P_I - P_V} - \sqrt{P_I + P_V} \right|}{\sqrt{P_I - P_V} + \sqrt{P_I + P_V}} \quad ,$$

whereas the angle χ that the major axis of the ellipse forms with the x-axis can be found through the equation

$$\tan(2\chi) = \frac{P_U}{P_Q} \quad .$$

The quantities P_I, P_Q, P_U, and P_V now introduced are not independent. Indeed they obey the relationship

$$P_I^2 = P_Q^2 + P_U^2 + P_V^2 \quad ,$$

and, varying their values, any kind of polarization ellipse can be described. Circular polarization is obtained by setting $P_Q = P_U = 0$, and one speaks about positive (or right-handed) circular polarization if $P_V = P_I$ and of negative (or left-handed) circular polarization if $P_V = -P_I$. In these cases the tip of the electric field vector describes a circle. On the other hand, linear polarization is obtained by setting $P_V = 0$. Now the tip of the electric vector oscillates along a segment whose inclination with respect to the x-axis is determined by the values of P_Q and P_U. In general, when none of the three quantities P_Q, P_U and P_V is zero, the tip of the electric vector describes an ellipse.

The description now given in terms of the polarization ellipse is however valid only for a plane, monochromatic wave which goes on indefinitely from $t = -\infty$ to $t = +\infty$. This is obviously a mathematical abstraction which, in general, has little to do with the physical world. A much more realistic description of a beam of radiation can be given only in terms of a statistical superposition of many wave-packets each having a limited extension in space and time. The beam thus loses its property of being monochromatic, becoming a quasi-monochromatic wave. Moreover, if the individual wave-packets do not share the same polarization properties, the polarization ellipse varies, statistically, in time. For such a beam of radiation it is then quite natural to generalize the previous definitions in the following form

$$\begin{aligned} P_I &= \langle E_1^2 + E_2^2 \rangle = \langle \mathcal{E}_1^* \mathcal{E}_1 \rangle + \langle \mathcal{E}_2^* \mathcal{E}_2 \rangle \quad , \\ P_Q &= \langle E_1^2 - E_2^2 \rangle = \langle \mathcal{E}_1^* \mathcal{E}_1 \rangle - \langle \mathcal{E}_2^* \mathcal{E}_2 \rangle \quad , \\ P_U &= \langle 2E_1 E_2 \cos(\phi_1 - \phi_2) \rangle = \langle \mathcal{E}_1^* \mathcal{E}_2 \rangle + \langle \mathcal{E}_2^* \mathcal{E}_1 \rangle \quad , \\ P_V &= \langle 2E_1 E_2 \sin(\phi_1 - \phi_2) \rangle = \mathrm{i}\left(\langle \mathcal{E}_1^* \mathcal{E}_2 \rangle - \langle \mathcal{E}_2^* \mathcal{E}_1 \rangle \right) \quad , \end{aligned} \tag{2.1}$$

where the symbol $\langle ... \rangle$ means an average over the statistical distribution of the wave-packets.

Through the new definitions one can indeed describe a much larger set of physical situations. In particular, being now

$$P_I^2 \geq P_Q^2 + P_U^2 + P_V^2 \quad ,$$

it is possible for a particular beam of radiation to have $P_Q = P_U = P_V = 0$. As it can be easily derived from the equations, this implies

$$\langle \mathcal{E}_1^* \mathcal{E}_1 \rangle = \langle \mathcal{E}_2^* \mathcal{E}_2 \rangle \, , \qquad \langle \mathcal{E}_1^* \mathcal{E}_2 \rangle = 0 \quad ,$$

which means that the electric field components along the x and y-axis are, in average, equal and uncorrelated. Such a beam is a beam of "natural" radiation and its description has been made possible by the "averaging" operation over the different wave packets. It is just this operation that has been introduced by Stokes in the description of polarized radiation and the quantities defined in Eqs.(2.1) are, apart from a dimensional factor needed to transform the square of an electric field into a specific intensity, just the Stokes parameters. The older descriptions of polarization, like the one used by Fresnel, did not take into account this averaging process and were then suitable to treat only totally polarized beams of radiation.

The description of polarization presented above involves suitable averages of the electric vibrations along two orthogonal axes, x and y, perpendicular to the direction of propagation. In practice, with the remarkable exception of radio-polarimetry, the electric field of a radiation beam cannot be measured directly, and it is then necessary to introduce some operational definitions in order to relate the polarization properties of a beam to actual measurements that can be performed on the beam itself. To reach this aim, it is convenient to refer to the concept of ideal polarizing filters, such as the *ideal polarizer* and the *ideal retarder*. These ideal devices are defined by specifying their action on the electric field components along two orthogonal axes perpendicular to the direction of propagation. For the ideal polarizer one has

$$\begin{pmatrix} \mathcal{E}'_{\mathrm{a}} \\ \mathcal{E}'_{\mathrm{b}} \end{pmatrix} = \mathrm{e}^{\mathrm{i}\psi} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{E}_{\mathrm{a}} \\ \mathcal{E}_{\mathrm{b}} \end{pmatrix} = \mathrm{e}^{\mathrm{i}\psi} \begin{pmatrix} \mathcal{E}_{\mathrm{a}} \\ 0 \end{pmatrix} \quad ,$$

where $\mathcal{E}_{\mathrm{a}}$ and $\mathcal{E}_{\mathrm{b}}$ are the components, at the entrance of the polarizer, of the electric field vector along the transmission axis and along the perpendicular axis, whereas $\mathcal{E}'_{\mathrm{a}}$ and $\mathcal{E}'_{\mathrm{b}}$ are the same components at the exit of the polarizer. As this equation shows, the electric field along the transmission axis is totally transmitted, whereas the transverse component is totally absorbed. The polarizer also manifests itself through a phase-factor, ψ, which is however completely inessential because it affects both components in the same way. For the ideal retarder, on the contrary, one has

$$\begin{pmatrix} \mathcal{E}'_{\mathrm{f}} \\ \mathcal{E}'_{\mathrm{s}} \end{pmatrix} = \mathrm{e}^{\mathrm{i}\psi} \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\delta} \end{pmatrix} \begin{pmatrix} \mathcal{E}_{\mathrm{f}} \\ \mathcal{E}_{\mathrm{s}} \end{pmatrix} = \mathrm{e}^{\mathrm{i}\psi} \begin{pmatrix} \mathcal{E}_{\mathrm{f}} \\ \mathrm{e}^{\mathrm{i}\delta} \mathcal{E}_{\mathrm{s}} \end{pmatrix} \quad ,$$

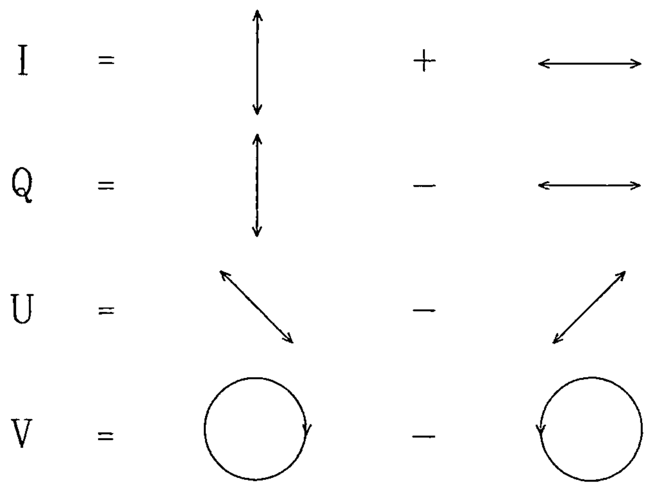

FIGURE 1. Pictorial representation of the Stokes parameters. The observer is supposed to face the radiation source.

where the notations are similar to those employed in the former equation and where the indices "f" and "s" stand, respectively, for the fast-axis and the slow-axis. The ideal retarder acts by introducing a supplementary phase factor, called *retardance* in the electric field component along the slow axis. If $\delta = \pi/2$, the retarder is also called a quarter-wave plate, if $\delta = \pi$, it is called a half-wave plate, and so on. It can be easily shown that the combination of a quarter-wave plate and a polarizer whose transmission axis is set at +45° (−45°) from the fast axis of the plate acts as a filter for positive (negative) circular polarization.

Through the ideal polarizing filters it is possible to give a simple, operational definition of the Stokes parameters of a beam of radiation. Consider a beam and a reference direction in the plane perpendicular to the beam. One starts by setting an ideal polarizer with its transmission axis along the reference direction and measures the intensity of the beam at the exit of the polarizer, thus obtaining the value $I_{0°}$. The same operation is repeated three times after rotating the polarizer (in the counterclockwise direction facing the source) of the angles 45°, 90°, and 135°, respectively, thus obtaining the values $I_{45°}$, $I_{90°}$, and $I_{135°}$. The ideal polarizer is then substituted by an ideal filter for positive circular polarization, the measured intensity at the exit of the filter being I_+, and by an ideal filter for negative circular polarization, the measured intensity being I_-. The operational definition of the four Stokes parameters, pictorially summarized in Fig. 1, is the following

$$I = I_{0°} + I_{90°} = I_{45°} + I_{135°} = I_+ + I_- \quad ,$$

$$Q = I_{0°} - I_{90°} \; , \qquad U = I_{45°} - I_{135°} \; , \qquad V = I_+ - I_- \quad .$$

By means of the properties of the ideal filters given previously, it is possible to relate the Stokes parameters with the quantities P_I, P_Q, P_U and P_V defined in Eqs.(2.1). When the reference direction introduced for the operational definition of the Stokes parameters coincides with the x-axis of the system introduced for the definition of the electric field components, one simply has

$$I = kP_I \;, \qquad Q = kP_Q \;, \qquad U = kP_U \;, \qquad V = kP_V \quad ,$$

where k is a dimensional constant whose precise value is often irrelevant because only the ratios Q/I, U/I and V/I are generally measured in practice.

3. Polarization and Optical Devices: Jones Calculus and Mueller Matrices

The ideal polarizer and the ideal retarder that we have considered above, are just two examples of optical devices for which a linear relationship of the form

$$\begin{pmatrix} \mathcal{E}'_1 \\ \mathcal{E}'_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \end{pmatrix} \tag{3.2}$$

can be established. In this equation, the unprimed components of the electric field refer to the beam at the entrance of the optical device, whereas the primed components refer to the exit beam. Moreover, a, b, c, and d are four complex quantities that define the physical characteristics of the optical device. This equation is the basis of the so-called *Jones calculus*, a particular formalism for treating polarization phenomena systematically introduced in the scientific literature by Jones in the early 1940s. The two-component vectors containing the electric field (in complex notations) are called Jones vectors, whereas the 2×2 matrix containing the properties of the optical device is called the Jones matrix. Obviously, for a train of N optical devices one can simply build up the Jones matrix of the train by considering the product of N individual 2×2 matrices:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a_N & b_N \\ c_N & d_N \end{pmatrix} \cdots\cdots \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \quad ,$$

where the first optical device encountered by the beam is characterized by the index 1, the second by the index 2, and so on (in other words, the ordering of the matrices in the r.h.s. is opposite to the ordering in which the optical devices are inserted in the beam).

The relationship between the electric field components of the entrance and exit beams given by Eq.(3.2) can be easily translated into a relationship between the Stokes parameters. Using the definition of the Stokes parameters, one obtains, after some algebra, an equation of the form

$$S' = \mathbf{M}\, S \quad , \tag{3.3}$$

where S is a 4-component vector constructed with the Stokes parameters of the entrance beam ($S^{\mathrm{T}} = (I, Q, U, V)$), S' has a similar meaning for the exit beam, and $\mathbf{M}$ is a 4×4 matrix given by

$$\frac{1}{2}\begin{pmatrix} a^*a + b^*b + c^*c + d^*d & a^*a - b^*b + c^*c - d^*d & 2\mathrm{Re}(a^*b + c^*d) & 2\mathrm{Im}(a^*b + c^*d) \\ a^*a + b^*b - c^*c - d^*d & a^*a - b^*b - c^*c + d^*d & 2\mathrm{Re}(a^*b - c^*d) & 2\mathrm{Im}(a^*b - c^*d) \\ 2\mathrm{Re}(a^*c + b^*d) & 2\mathrm{Re}(a^*c - b^*d) & 2\mathrm{Re}(a^*d + b^*c) & 2\mathrm{Im}(a^*d - b^*c) \\ -2\mathrm{Im}(a^*c + b^*d) & -2\mathrm{Im}(a^*c - b^*d) & -2\mathrm{Im}(a^*d + b^*c) & 2\mathrm{Re}(a^*d - b^*c) \end{pmatrix}$$

A 4×4 matrix as the one here introduced is usually referred to as a Mueller matrix. Such a matrix is made, in general, of 16 independent elements and the expression that we have derived above (which depends indeed on only 7 quantities –the real and imaginary parts of the 4 elements a, b, c, and d of the Jones matrix, minus an irrelevant phase that can be factorized in the same matrix) is a particular case of a Mueller matrix. In the following, we will refer to this particular case as the Jones-Mueller matrix.

The peculiarity of a Jones-Mueller matrix is contained in a subtle mathematical property which we state here without proof. If the determinant of the Jones matrix is non-zero, that is if

$$D = ad - bc \neq 0 \quad ,$$

then it follows that

$$|D|^2 \mathbf{M}^{-1} = \mathbf{X}\mathbf{M}^{\mathrm{T}}\mathbf{X} \quad ,$$

where $\mathbf{X}$ is the diagonal matrix defined by

$$\mathbf{X} = \mathbf{X}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad . \tag{3.4}$$

Through this mathematical property it is possible to show an interesting result for the polarization properties of the entrance and exit beams connected by a Jones-Mueller matrix. Defining

$$\mathcal{P} = I^2 - Q^2 - U^2 - V^2 = S^{\mathrm{T}}\mathbf{X}S \;\; ; \qquad \mathcal{P}' = I'^2 - Q'^2 - U'^2 - V'^2 = S'^{\mathrm{T}}\mathbf{X}S' \quad , \tag{3.5}$$

one has, with easy transformations

$$\mathcal{P}' = S'^{\mathrm{T}}\mathbf{X}S' = S^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{X}\mathbf{M}S = |D|^2 S^{\mathrm{T}}\mathbf{X}\mathbf{M}^{-1}\mathbf{M}S = |D|^2\mathcal{P} \quad .$$

The equation connecting the first and last terms of this chain of equalities, which can be proved to be valid also in the case where $|D|^2 = 0$, shows that: a) if $\mathcal{P} \geq 0$, also $\mathcal{P}' \geq 0$; b) if $\mathcal{P} = 0$, then $\mathcal{P}' = 0$. Property a) means that a Jones-Mueller matrix is always a physical (or bona-fide) Mueller matrix, in the sense that it transforms physical polarization states ($\mathcal{P} \geq 0$) in physical polarization states ($\mathcal{P}' \geq 0$). Property b) shows that a totally polarized beam is always transformed by a Jones-Mueller matrix into another totally polarized beam. In other words a Jones-Mueller matrix is unable of describing depolarizing mechanisms and this clearly shows the limitations of the Jones calculus for handling a large variety of polarization phenomena. As an example, consider the case of an ideal depolarizer. The corresponding Mueller matrix is obviously given by an expression of the form

$$\mathbf{M}_{\text{ideal depolarizer}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad ,$$

It can be easily proved that it is impossible to find a set of values for the quantities a, b, c, and d, such that, when substituted in the expression for the Jones-Mueller matrix, are capable of reproducing the Mueller matrix of the ideal polarizer.

Mueller matrices have a large variety of applications in physics and, more particularly, in astronomy. In many cases, one can even define the Mueller matrix of a telescope by analyzing the properties of each of its optical devices and then deducing the resulting matrix as the product of the matrices of each device. The "train property" outlined for the Jones matrices is obviously valid for the Mueller matrices too, so that one has, with evident notations

$$\mathbf{M} = \mathbf{M}_N ... \mathbf{M}_2 \mathbf{M}_1 \quad .$$

An important problem about Mueller matrices, that often arises when one is trying to deduce the Mueller matrix of an optical device (or a combination of several optical devices) by means of experiments, is the following: given a 4×4 real matrix whose 16 elements are to be considered as quantities affected by experimental errors, is it a physical (or bona-fide) Mueller matrix, or not? This problem has been solved quite recently by means of a mathematical algorithm directly implemented in a code (Landi

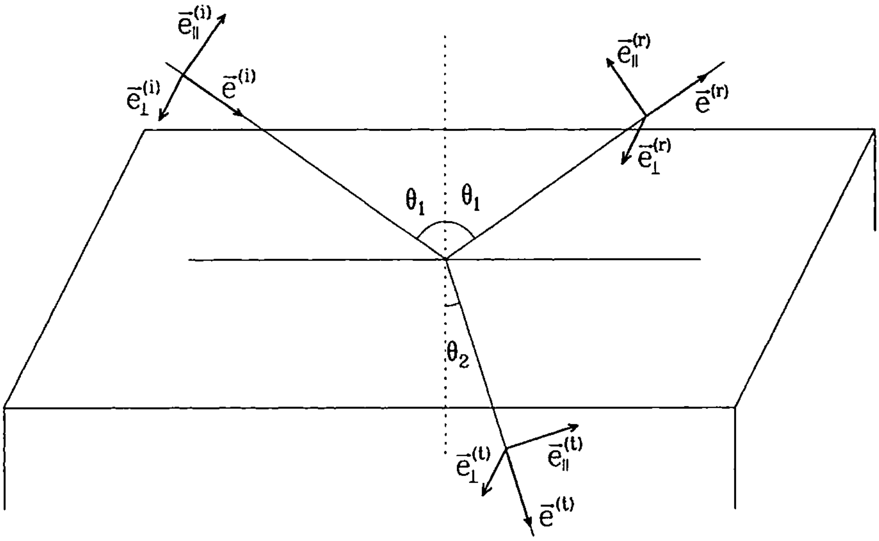

FIGURE 2. Reflection and refraction at the separation of two dielectrics.

Degl'Innocenti & del Toro Iniesta, 1998). As a curiosity we can mention the fact that the number of physical Mueller matrices is a very tiny fraction of all the 4×4 matrices that can be constructed. More precisely, calculations performed via a Montecarlo technique show that out of 10^6 matrices generated by setting the element M_{11} to 1 and all the other elements to random numbers bound in the interval $(-1, 1)$, only approximately two matrices turn out to be bona-fide Mueller matrices.

4. The Fresnel Equations

The simplest and commonest physical phenomenon where polarization processes enter into play is the ordinary reflection of a pencil of radiation on the surface of a dielectric medium. This phenomenon, which is generally accompanied by the related phenomenon of refraction, is described by the so-called Fresnel equations that can be derived as a direct consequence of the Maxwell equations. Referring to Fig. 2, we denote by n_1 and n_2 the index of refraction of the two media by θ_1 the angle of incidence (which is equal to the angle of reflection) and by θ_2 the angle of refraction. Considering, for the time being, the simplest case where both media (1 and 2) are dielectrics (which implies that n_1 and n_2 are real), and supposing $n_1 \leq n_2$, the angles θ_1 and θ_2 are connected by the usual Snell's law

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad . \tag{4.6}$$

The incident, the reflected, and the refracted ray all lie in the same plane which also contain the normal to the surface of separation between the two media (the so-called *incidence plane*). For each ray, a right-handed reference frame is introduced, with the third axis directed along the ray, the first axis lying in the plane of incidence, and the second axis being directed perpendicularly to the plane of incidence (and being then parallel to the surface of separation of the two media). The unit vectors are denoted respectively as $(\vec{e}_{\parallel}^{\,(i)}, \vec{e}_{\perp}^{\,(i)}, \vec{e}^{\,(i)})$ for the incident beam, $(\vec{e}_{\parallel}^{\,(r)}, \vec{e}_{\perp}^{\,(r)}, \vec{e}^{\,(r)})$ for the reflected beam, and $(\vec{e}_{\parallel}^{\,(t)}, \vec{e}_{\perp}^{\,(t)}, \vec{e}^{\,(t)})$ for the refracted (or transmitted) beam. Using similar notations for denoting the electric field components along the different axes, the laws of Fresnel are condensed by the following equations

$$\begin{pmatrix} \mathcal{E}_{\parallel}^{(r)} \\ \mathcal{E}_{\perp}^{(r)} \end{pmatrix} = \begin{pmatrix} r_{\parallel} & 0 \\ 0 & r_{\perp} \end{pmatrix} \begin{pmatrix} \mathcal{E}_{\parallel}^{(i)} \\ \mathcal{E}_{\perp}^{(i)} \end{pmatrix} , \qquad \begin{pmatrix} \mathcal{E}_{\parallel}^{(t)} \\ \mathcal{E}_{\perp}^{(t)} \end{pmatrix} = \begin{pmatrix} t_{\parallel} & 0 \\ 0 & t_{\perp} \end{pmatrix} \begin{pmatrix} \mathcal{E}_{\parallel}^{(i)} \\ \mathcal{E}_{\perp}^{(i)} \end{pmatrix} ,$$

where

$$r_{\parallel} = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2} , \qquad r_{\perp} = \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2} ,$$
$$t_{\parallel} = \frac{2 n_1 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2} , \qquad t_{\perp} = \frac{2 n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2} . \tag{4.7}$$

Since the equations now derived are in the form of "Jones equations" (cfr. Eq.(3.2)), it is easy to find the Mueller matrices corresponding to reflection and to refraction (or transmission). Taking into account Eq.(3.3) and choosing for each of the three rays the reference direction along $\vec{e}_{\parallel}$, we find for reflection

$$\mathbf{M}_{\text{reflection}} = \frac{1}{2} \begin{pmatrix} |r_{\parallel}|^2 + |r_{\perp}|^2 & |r_{\parallel}|^2 - |r_{\perp}|^2 & 0 & 0 \\ |r_{\parallel}|^2 - |r_{\perp}|^2 & |r_{\parallel}|^2 + |r_{\perp}|^2 & 0 & 0 \\ 0 & 0 & 2\text{Re}(r_{\parallel}^* r_{\perp}) & 2\text{Im}(r_{\parallel}^* r_{\perp}) \\ 0 & 0 & -2\text{Im}(r_{\parallel}^* r_{\perp}) & 2\text{Re}(r_{\parallel}^* r_{\perp}) \end{pmatrix} ,$$

and, for transmission,

$$\mathbf{M}_{\text{transmission}} = \frac{n_2 \cos\theta_2}{n_1 \cos\theta_1} \frac{1}{2} \begin{pmatrix} |t_{\parallel}|^2 + |t_{\perp}|^2 & |t_{\parallel}|^2 - |t_{\perp}|^2 & 0 & 0 \\ |t_{\parallel}|^2 - |t_{\perp}|^2 & |t_{\parallel}|^2 + |t_{\perp}|^2 & 0 & 0 \\ 0 & 0 & 2\text{Re}(t_{\parallel}^* t_{\perp}) & 2\text{Im}(t_{\parallel}^* t_{\perp}) \\ 0 & 0 & -2\text{Im}(t_{\parallel}^* t_{\perp}) & 2\text{Re}(t_{\parallel}^* t_{\perp}) \end{pmatrix} .$$

In this last equation, a supplementary factor $n_2 \cos\theta_2 / (n_1 \cos\theta_1)$ has been introduced in front of the matrix to account for the fact that the energy that is contained, in the incident beam, within the infinitesimal angle $\text{d}\theta_1$, is contained, after refraction, within the different infinitesimal angle $\text{d}\theta_2$. On the other hand, from Snell's law (Eq.(4.6)), one has

$$\frac{\text{d}\theta_1}{\text{d}\theta_2} = \frac{n_2 \cos\theta_2}{n_1 \cos\theta_1} .$$

An important property of the Fresnel equations is the fact that they are capable of describing, besides the phenomenon of reflection and refraction at the surface of two dielectrics, with the radiation propagating from the less refracting to the more refracting medium, also the inverse phenomenon where a pencil of radiation is propagating from a more refractive medium to a less refracting medium, and also the phenomenon of reflection on the surface of a metal. For treating these two supplementary cases, which require some further conventions, it is convenient to rewrite Eqs.(4.7) in the equivalent form

$$r_{\parallel} = \frac{n_2^2 u_1 - n_1^2 u_2}{n_2^2 u_1 + n_1^2 u_2} , \qquad r_{\perp} = \frac{u_1 - u_2}{u_1 + u_2} ,$$
$$t_{\parallel} = \frac{2 n_1 n_2 u_1}{n_2^2 u_1 + n_1^2 u_2} , \qquad t_{\perp} = \frac{2 u_1}{u_1 + u_2} , \tag{4.8}$$

where

$$u_1 = n_1 \cos\theta_1 , \qquad u_2 = n_2 \cos\theta_2 .$$

Consider first the case of two dielectrics with $n_1 > n_2$. A direct application of Snell's