THE THEORY OF POLARIZATION PHENOMENA

BY

B. A. ROBSON

CLARENDON PRESS · OXFORD
1974

Oxford University Press, Ely House, London W.1

GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
CAPE TOWN IBADAN NAIROBI DAR ES SALAAM LUSAKA ADDIS ABABA
DELHI BOMBAY CALCUTTA MADRAS KARACHI LAHORE DACCA
KUALA LUMPUR SINGAPORE HONG KONG TOKYO

ISBN 0 19 851453 0

PRINTED IN GREAT BRITAIN
BY J. W. ARROWSMITH LTD, BRISTOL BS3 2NT

OXFORD STUDIES IN PHYSICS

PREFACE

THE purpose of this book is to provide a detailed but simple development of the general formalism required to describe the polarization of particles of arbitrary spin and their decay into or interaction with other particles. The book evolved out of two series of lectures given at the Australian National University, Canberra, mainly to post-graduate students and research workers in the field of nuclear physics. There appeared to be a definite need for a book devoted entirely to the theory of polarization phenomena. While many books of a more general nature contain some discussion of the polarization of particles, in most cases the treatment is too superficial and inadequate. On the other hand the specialist papers and review articles tend to lack sufficient introduction to the subject for the average research worker to grasp the underlying physics of the often quite complicated theoretical formalism.

In this book I have tried to show the unity and logical development of the subject from the initial discovery of the polarization of light. In particular I have given considerable emphasis to the method developed by Stokes, Soleillet, Perrin, and Mueller in classical optics now known as the Mueller calculus. This approach, which offers considerable simplicity and clarity both for designing polarization experiments and for understanding the resultant polarization measurements, has been almost completely ignored outside of optics. I hope that the present book will rectify this situation.

The book is intended to be read from the beginning to the end. It has been written at the level of a graduate physics course and assumes a basic knowledge of quantum theory, scattering theory, and matrix algebra. It should be useful to research workers in all branches of physics (optics, atomic, nuclear, particle, etc.) who study polarization effects.

The book is essentially pure theory with no experimental results being discussed. Thus no attempt has been made to consider all polarization phenomena; only a few simple examples (mostly from my own field of nuclear physics) are presented in order to illustrate the theoretical formalism.

As far as possible I have adopted the Madison convention for specifying polarization quantities. Unfortunately, this convention is not consistent with the logical definition of the analysing powers for the scattering of a polarized incident beam in the spherical tensor representation. For this reason I have denoted the Mueller matrix by Z rather than T. The Madison convention does not include quantities such as polarization transfer and spin-correlation coefficients. I hope that the notation and definition of these quantities adopted in this book will find ready acceptance by research workers.

I have taken particular care to differentiate between quantities which specify the polarization of a beam of particles and quantities which specify the interaction of such particles with a target. Thus I have introduced the quantity 'vector scattering parameter' for the elastic scattering of spin-$\frac{1}{2}$ particles by a spinless target. Although this quantity is sometimes numerically equal to the vector polarization of the scattered beam, it is confusing to identify the two quantities too closely.

I wish to express my gratitude to Dr. N. Berovic who read almost all of the manuscript and gave me very valuable criticism and advice. I am also indebted to many colleagues for discussions and comments on the manuscript and Mrs. I. Kinchin for preparing most of the final manuscript in such a cheerful and competent manner.

I am grateful to Profs. W. E. Burcham and F. Beck for hospitality at the University of Birmingham and the Institut für Kernphysik, Darmstadt, respectively, where considerable portions of the book were written. I also thank Prof. K. J. LeCouteur for encouragement to commence such a project.

B.A.R.

The Australian National University, Canberra
July 1974

CONTENTS

INTRODUCTION

THE observation by Bartholinus in 1669 of the double refraction of light by calcite led to the discovery of the *polarization* of light by Huygens *c.* 1690. Huygens found that light which had passed through a piece of calcite behaved differently from ordinary light. However, although he was able to describe the phenomenon of double refraction in terms of his wave construction, he was unable to account for the polarization of light. This was not achieved until *c.* 1817 when Young suggested that light waves are transverse rather than longitudinal vibrations (Fresnel claimed to have mentioned this possibility to Ampère in 1816). In 1824 Fresnel showed that light waves are exclusively transverse and the resultant transverse vector theory constituted the first theory of polarized light. This was eventually superseded by the more general electromagnetic theory of Maxwell in 1864.

The term *polarization* was first used by Malus in 1810 when describing the production of polarized light by reflection and was derived from the word 'polarity' employed much earlier to describe the two-sidedness or two-fold nature of magnetic poles. Malus, while keeping an open mind on the wave-versus-corpuscular theories of light, employed the latter model and considered that the polarization of light was connected with the polarity of the corpuscles. In this sense the word polarization is a misnomer but the term has such a long history that there can be no question of a replacement for it. Unfortunately, polarization has also been used to describe other effects, e.g. in Maxwell's displacement vector $\mathbf{D} = \mathbf{E}+4\pi\mathbf{P}$. Here $\mathbf{E}$ is the electric vector and the polarization vector $\mathbf{P}$ is a measure of the mean polarizability of a dielectric medium. We do not consider such phenomena in this treatise.

The next polarization phenomena to be observed and described were essentially the Zeeman effect and the doublet spectral lines of the alkali elements. In 1896 Zeeman discovered that certain spectral lines are split into a number of components on the application of an external magnetic field. The classical theory of Lorentz indicated the splitting of lines into three components (the normal Zeeman effect) and indeed in some cases is able to account for the measurements. However, in many instances there occur more than three components—the so-called anomalous Zeeman effect. For the explanation of this latter phenomenon and the alkali doublet structure, it was necessary to assume that the electron possesses an intrinsic angular momentum called spin, which by analogy with the quantization of orbital angular momentum gives rise to just two basic states, i.e. electrons are spin-$\frac{1}{2}$ particles. The spin of the electron gives rise to a small splitting of the majority of alkali atomic energy levels and the corresponding occurrence of doublet spectral lines. Associated with the spin angular momentum is a magnetic moment which accounts for the anomalous Zeeman effect.

Since 1925, when the concepts of electron spin and magnetic moment were published, each particle or system of particles (e.g. deuteron) is considered to possess a spin s which has a unique value from the set $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. Thus pions and alpha particles have $s = 0$, electrons and protons have $s = \frac{1}{2}$, photons and deuterons have $s = 1$, etc. All particles with $s > 0$ exhibit polarization phenomena, i.e. effects which arise as a direct consequence of their intrinsic spin. This book deals only with the description of such polarization phenomena.

For several reasons we discuss polarized light first (Chapter 1). Since the polarization of light was studied for over two hundred years before any other polarization phenomenon, much of the terminology of polarization theory is derived from optics. Secondly, polarized light offers simple and convenient examples with which to introduce the various formalisms. Moreover, this can be done using a classical approach which at a later stage and as a separate step may be simply re-interpreted in a quantum-theory treatment (Chapter 2). The two main approaches now employed in optics for describing the interaction of polarized light with optical devices are the Jones and Mueller calculi, which were invented in the early 1940s. Both these matrix methods are conveniently introduced by considering the passage of polarized light through two simple optical instruments, the 'stopped' calcite crystal (or its equivalent the Nicol prism) and the quarter-wave plate, which have comparatively trivial matrix representations. Furthermore, from their study of classical optics, many readers undoubtedly will be familiar with the properties of these devices as well as the different forms of polarized light. Finally, the photon concept played a leading role in the development of quantum theory and presents a convenient stepping stone for the transition to a general description of polarized particles.

The two calculi employed in classical optics have their counterparts in the description of the polarization of particles of arbitrary spin and their decay or interaction with other particles. Indeed rather earlier than 1940, the equivalent of the Jones calculus, namely, the use of spin wave functions and scattering (or reaction) matrices was introduced into particle physics by Pauli in 1927 and Wheeler in 1937, respectively. The extension of the Jones calculus to include partially polarized light requires the use of the density matrix proposed by von Neumann in 1927. On the other hand, the Mueller method has scarcely been employed for particles. However, it is the author's belief that this complementary approach offers considerable simplicity compared with the usual description based upon density and reaction matrices. Both methods and the relationship between them are discussed in detail (Chapters 3–6).

In Chapters 3 and 4 we discuss in depth the complete specification of the polarization of spin-$\frac{1}{2}$ and spin-1 particles, respectively. Emphasis is also given to spin-1 particles since they exhibit fundamental differences from the simplest case of spin-$\frac{1}{2}$ particles which are typical of higher spin particles. Our over-all approach is to proceed by analogy and induction from the simpler to the more complicated phenomena rather than to commence with a general formalism. This necessitates some repetition, but the author believes that this gradual development of the theory is more understandable. Thus only the simplest reaction, elastic scattering from spinless targets, is considered at this stage. The general forms of the elastic scattering matrices under certain invariance requirements (e.g. parity conservation) and the corresponding numbers of independent observable quantities are discussed. Throughout we employ a notation to denote reference axes which is becoming more essential as the trend now is to refer initial and final spin states to different coordinate frames. We use a set of 'standard' axes which correspond to the helicity coordinate systems of Ohlsen (1972).

In Chapter 5 we present the general non-relativistic formalism for particles of arbitrary spin interacting with targets with spin. Both non-elastic reactions and processes involving identical particles are discussed. The formalism is then extended to the emission and absorption of electromagnetic radiation (Chapter 6). The concepts of decay and absorption matrices are introduced and it is believed that the treatment of angular correlations presented here is both novel and simpler than previous descriptions.

Finally we include a short chapter on the treatment of relativistic particles. We follow the approach of Chou and Shirokov (1958) which represents (at least to the author) the only correct explicit treatment of the spin precession for an arbitrary proper Lorentz transformation. We indicate how this relativistic rotation of the spin may be easily incorporated into the non-relativistic formalism. We also show the relation of our formalism to the helicity representation of Jacob and Wick (1959). For various reasons we have not adopted the helicity formalism from the outset but in the case of particles with mass this is shown to be no disadvantage for the standard axes chosen. The case of massless particles is also considered.

1

POLARIZED LIGHT

1.1. Calcite crystal experiment

CALCITE is a rhombohedral crystalline form of calcium carbonate in which the double refraction of light is very strikingly exhibited. In a perfectly formed crystal (Fig. 1.1) the rhombohedron is bounded by six similar parallelograms the obtuse angles of which are about 101° 55′. The solid angles at the corners A and G are contained by three obtuse angles while the remainder are bounded by one obtuse and two acute angles. The plane ACGE and the line AG are called the principal plane and the principal axis respectively. A calcite crystal is said to be uniaxial since there is one special direction called the optic axis in which only single refraction occurs. For calcite the optic axis coincides with the direction of the principal symmetry axis AG.

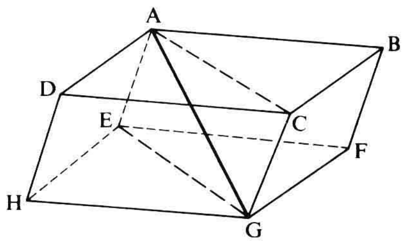

FIG. 1.1. Perfectly formed calcite crystal with principal plane ACGE and principal (optic) axis AG.

When a beam of ordinary light is passed through a slab of calcite crystal each ray is generally divided into two, an ordinary (O) ray which obeys the usual law of refraction and a so-called extraordinary (E) ray which does not. This phenomenon is undoubtedly related to the crystalline structure of calcite since substances such as glass which have an irregular structure do not exhibit such an effect. For convenience let us assume that the incident direction is not too close to the optic axis and is both normal to the crystal face and in a principal plane, i.e. in a direction parallel to the principal plane (Fig. 1.2). The O-ray passes straight through while the E-ray is deflected along a principal plane and emerges parallel to the incident ray. Thus a rotation of slab 1 about the incident direction causes an equal rotation of the E-ray about the same direction. When these two rays strike the face of slab 2 and provided both crystals have identical orientation of their optic axes, the O-ray continues undeflected while the E-ray is further displaced. This shows that light which has traversed a calcite crystal is different from ordinary light.

The light is said to be *polarized* and *unpolarized* respectively. Furthermore, the O- and E-rays are different so that light can exist in at least two different polarized states. However, it is necessary to be careful here; one should not visualize light as an incoherent mixture of two kinds X and Y, which are separated by the calcite slab so that the O-ray consists of type X and the E-ray of type Y. If this were so, it would not be possible to obtain a further separation in the second crystal as does in fact occur if the orientations of the optic axes are different (Fig. 1.3). In this case the calcite slab transforms the set of states $[X, Y]$ into other polarization states $[1, 2, 3, 4]$ depending upon its orientation α.

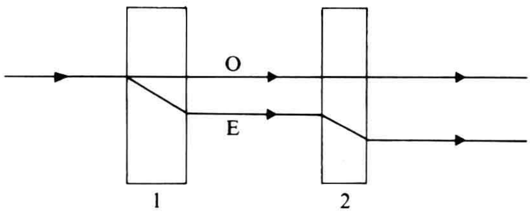

FIG. 1.2. Double refraction of ordinary (unpolarized) light by two successive calcite crystals having the same orientation of optic axes. Both the ordinary (O) and extraordinary (E) rays are polarized. All rays are in the plane of the page.

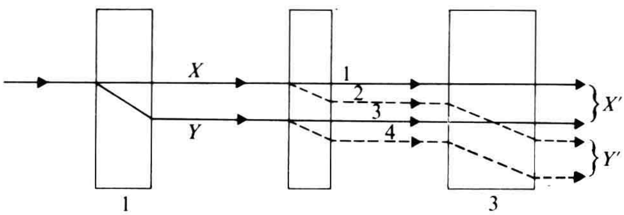

FIG. 1.3. Double refraction of ordinary (unpolarized) light by three successive calcite crystals with crystals 2 and 3 having the same orientation of optic axes. The X and Y rays each separate into two components in crystal 2. Rays 1 and 3 (and likewise rays 2 and 4) behave similarly on passing crystal 3. Rays 2 and 4 represented by broken lines are not in the plane of the page.

If the four rays are passed through a third calcite crystal which has its optic axis in the same direction as the second crystal, it is found that rays 1 and 3 behave similarly and also that rays 2 and 4 are alike. Thus there are only two types of polarization states again. Indeed, the interaction of light with all types of optical devices can be satisfactorily described in terms of just two states of polarization. It is therefore possible to write the effect of a calcite crystal upon polarized light as

$$[X', Y'] = T(\alpha)[X, Y] \tag{1.1}$$

which means that the polarization states X, Y are transformed by the operator $T(\alpha)$, which describes in some manner the orientation and action of the crystal, into the polarization states X', Y'. This form of relationship immediately suggests that light cannot be a scalar quantity; it must have components. The phenomenon of polarization shows that a satisfactory theory of light has to be vectorial in character, i.e. be concerned with directions. It should be noted that if one starts with ordinary light two calcite crystals are required to observe the polarization effect, the first to polarize the light and the second to analyse it. One speaks of a *polarizer* and an *analyser* respectively and of course a polarizer can act as an analyser and vice versa.

1.2. Electromagnetic theory of light

The electromagnetic theory represents light in free space as transverse vibrations of electric and magnetic fields **E**, **H** which are mutually perpendicular and in phase. For the purposes of the discussion here, it is only necessary to consider one of the vectors, say **E**, since the vectors are related to each other in terms of simple constants so that **H** can be derived from **E**.

The two rays of light transmitted by a calcite crystal are both *linearly polarized*, i.e. the electric vectors representing the rays have fixed directions so that each vibration takes place in a single plane containing the direction of propagation. Moreover, these two directions are at right angles to one another. It has been found that the polarization of light can be understood in terms of the superposition of two such rays; a given polarized ray of light may be represented as the resultant of two disturbances, one with the **E**-vector in the xz-plane and the other with the **E**-vector in the yz-plane and both travelling along the z-axis. We can write

$$\mathbf{E} = E_x\mathbf{e}_x + E_y\mathbf{e}_y, \tag{1.2}$$

where

$$E_x = a\cos(kz-\omega t) = a\cos\phi, \tag{1.3}$$

and

$$E_y = b\cos(kz-\omega t+\delta) = b\cos(\phi+\delta). \tag{1.4}$$

Here $\mathbf{e}_x$, $\mathbf{e}_y$ are unit vectors along the x-, y-axes, ω is the angular frequency, and k is the wave number. The two component vibrations have the same frequency and velocity of propagation but their amplitudes differ and there is a permanent phase difference δ. In general, the tip of the **E**-vector will appear to trace out an ellipse (the polarization ellipse) when viewed along the direction in which the light propagates, and consequently a polarized ray of light is said to be *elliptically polarized*. If the **E**-vector rotates around the ellipse in a clockwise (anti-clockwise) direction when viewed by an observer who receives the beam of light, the ray is said to be right-handed (left-handed) elliptically polarized light. There are two special cases: (1) if $a = b$ and $\delta = \frac{1}{2}m\pi$ ($m = \pm 1, \pm 3, \pm 5, \ldots$) the ellipse becomes a circle and the light is called right-handed and left-handed *circularly polarized* light respectively; (2) if $\delta = m\pi$ $(m = 0, \pm 1, \pm 2, \ldots)$ the ellipse degenerates to a straight line (the polarization line) and the light *is linearly polarized.*

1.3. Stokes parameters

To specify the polarization ellipse we require three independent quantities, e.g. the amplitudes a and b and the phase difference δ. For practical purposes it is convenient to characterize the state of polarization by certain parameters which are all of the same physical dimensions and which were introduced by Stokes (1852). For a plane monochromatic wave the *Stokes parameters* are the four quantities:

$$I = a^2+b^2, \qquad P_1 = a^2-b^2, \qquad P_2 = 2ab\cos\delta, \qquad P_3 = 2ab\sin\delta. \quad (1.5)$$

Only three of these quantities are independent since

$$I^2 = P_1^2+P_2^2+P_3^2. \quad (1.6)$$

The Stokes parameters are useful because they (1) can be determined by simple experiments, (2) allow treatment of unpolarized and partially polarized beams of light, and (3) may be re-interpreted in terms of the quantum theory of light. The parameter I is a measure of the *time-averaged intensity* of the wave, the average being taken over a period long enough to allow adequate measurement but very long indeed compared with the natural period $\omega \sim 10^{-15}\,\mathrm{s}^{-1}$ of the wave. The parameters P_1, P_2, and P_3 can be expressed in terms of quantities which describe the polarization ellipse, ψ the angle between the major axis and the x-axis, and $\chi = \tan^{-1}(\pm B/A)$, $-\frac{1}{4}\pi \leqslant \chi \leqslant \frac{1}{4}\pi$, where A and B are the lengths of the major and minor semi-axes (Fig. 1.4). The quantity $\tan\chi$ is defined to have the same sign as $\sin\delta$ and is positive

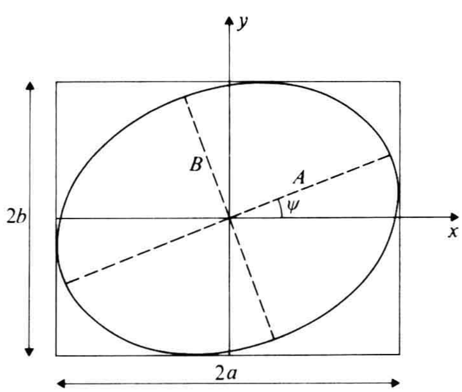

FIG. 1.4. Polarization ellipse for electric vector $\mathbf{E} = a\cos\phi\mathbf{e}_x + b\cos(\phi+\delta)\mathbf{e}_y$. The major and minor semi-axes have lengths A and B, respectively; ψ is the angle between the major axis and the x-axis.

(negative) for right-handed (left-handed) polarization. We have

$$P_1 = I \cos 2\chi \cos 2\psi, \qquad (1.7a)$$

$$P_2 = I \cos 2\chi \sin 2\psi, \qquad (1.7b)$$

$$P_3 = I \sin 2\chi. \qquad (1.7c)$$

These relationships bear a close resemblance to the formulae for the three components of a vector expressed in spherical coordinates and indicate a simple geometrical representation of all the different states of polarization. This representation is known as the *Poincaré sphere* (Poincaré 1892).

The Poincaré sphere Σ is a sphere of radius I (Fig. 1.5) such that any point P on the surface having spherical angular coordinates $(\frac{1}{2}\pi - 2\chi)$ and 2ψ represents one and only one state of polarization of a plane monochromatic wave. The reverse is also true; each point on the surface Σ uniquely defines one state of polarization. In particular:

(1) all points 'north' ('south') of the equator represent states of right-handed (left-handed) polarization;
(2) the north (south) pole represents right-handed (left-handed) circular polarization;
(3) the equator represents all states of linear polarization.

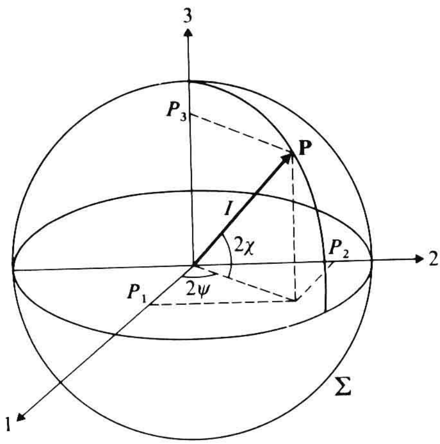

FIG. 1.5. Poincaré sphere (Σ) representation of all polarization states of a plane monochromatic wave. The polarization vector $\mathbf{P} = (P_1, P_2, P_3)$ has length I and polar angles $(\frac{1}{2}\pi - 2\chi)$, 2ψ.

1.4. Modern theories of polarized light

The description of the interaction of polarized light with several optical devices using conventional algebraic and trigonometric methods is a very difficult and cumbersome process. Two 'modern' methods which greatly simplify the description are the *Jones* and *Mueller calculi*. Both methods use