Eur. J. Phys. 18 (1997) 256–262. Printed in the UK PII: S0143-0807(97)79468-6

An elementary approach to spinors

F Gori†, S Vicalvi‡, M Santarsiero†, F Frezza§, G Schettini§, S Ambrosini‖ and R Borghi‡

† Dipartimento di Fisica, Università di Roma Tre, Via della Vasca Navale, 84, 00146 Roma, Italy
‡ Dipartimento di Fisica, Università 'La Sapienza', Piazzale A Moro, 2, 00185 Roma, Italy
§ Dipartimento di Ingegneria Elettronica, Università 'La Sapienza', Via Eudossiana, 18, 00184 Roma, Italy
‖ Dipartimento di Energetica, Università dell'Aquila, Loc. Monteluco di Roio, 67040 Roio Poggio-AQ, Italy

Received 12 November 1996

Abstract. We give a simple rule to associate a pair of complex numbers to a spinor. This association rule is suggested by an analysis of the link between Jones vectors and the Poincaré sphere, which are tools commonly used in the description of polarized light, and allows the basic properties of spinors, such as the effect of rotations, to be derived in a simple way. In our treatment, we give a purely mathematical description of spinors, without using the physical properties of the spin, so that the theory of spinors can be introduced independently from its application to physics.

Riassunto. In questo lavoro forniamo una semplice regola per associare una coppia di numeri complessi ad uno spinore. Tale regola è suggerita da un'analisi del legame che sussiste tra i vettori di Jones e la sfera di Poincaré, strumenti comunemente usati nella descrizione della luce polarizzata. Essa consente di derivare in modo semplice le proprietà di base degli spinori, come ad esempio l'effetto delle rotazioni. Nella nostra analisi, gli spinori sono descritti da un punto di vista puramente matematico, senza fare uso delle proprietà fisiche dello spin; in tal modo la teoria degli spinori può essere sviluppata indipendentemente dalla sua applicazione alla fisica.

1. Introduction

The introduction of spin and spinors is one of the most challenging tasks in an elementary course on quantum mechanics. Even in the simplest case of spin-$\frac{1}{2}$ certain fundamental aspects of this topic, such as the effects of rotation on spinors, turn out to be difficult to explain. Very often one makes recourse both to intuition and to physical constraints in order to derive the transformation rules. In some cases, it seems that arbitrary choices are made. As a result the student may be left with the unsatisfactory feeling that one has to figure out what the pertinent laws are by a sort of trial and error procedure. While somehow mirroring the historical development of the subject, such a continuous interplay between the physical system and the mathematical model may give the impression that the latter could not be constructed, at least in simple terms, without making appeal to the physical situation.

In order to avoid this state of affairs a clear cut distinction between construction of the mathematical tool (elementary spinor theory) and its applications to physics is useful. Concerning this, some purely mathematical approaches are present in the literature, making use of geometrical structures, such as the isotropic vector [1], the stereographic projection [2] and the flag picture [3].

In this paper we suggest an approach in which spinors (of the simplest type) are introduced as a tool for representing oriented lines in real space using a pair of complex numbers. The way to do this is suggested by an analysis of the link between two of the most popular methods for describing polarized light, namely, Jones vectors and the Poincaré sphere. Once the convention for associating spinors to unit vectors of real space is established, the basic facts of spinor calculus can be derived in a logical and straightforward way. This applies to transformation rules under arbitrary rotations, Pauli matrices and their commutation relations. In this approach the mathematics of spinors is introduced and developed without making use of spin properties (except for the starting motivation). In this way the application of such a tool to the description of spin appears as a logically separated step. In this paper, however, we shall not examine in detail such an application.

2. Jones vectors and the Poincaré sphere

The aim of this section is to underline certain aspects of Jones vectors and the Poincaré sphere that are useful for our work. Only an elementary knowledge of these two tools for representing polarized light is required. Let us

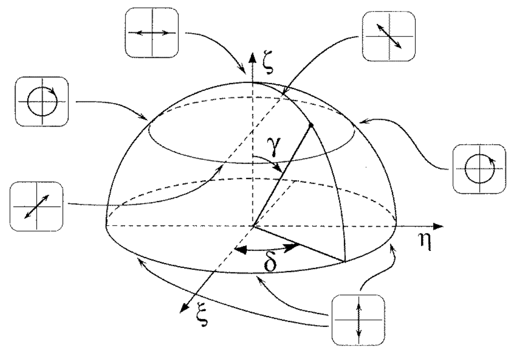

Figure 1. Light polarization states mapped on a hemisphere.

recall that a polarized monochromatic plane wave can be specified by the Jones vector [4, 5].

$$\begin{pmatrix} E_{0x}\,\mathrm{e}^{\mathrm{i}\varphi_x} \\ E_{0y}\,\mathrm{e}^{\mathrm{i}\varphi_y} \end{pmatrix}, \qquad (1)$$

where E_{0x} and E_{0y} are the amplitudes of the x- and y-components of the electric field while φ_x and φ_y are the corresponding initial phases. The polarization state remains unchanged both on multiplying E_{0x} and E_{0y} by a common factor and on adding a common angle to φ_x and φ_y. As far as the polarization state is concerned, we can thus reduce the Jones vector to the symmetric form

$$\begin{pmatrix} A_x \mathrm{e}^{-\mathrm{i}\delta/2} \\ A_y\,\mathrm{e}^{\mathrm{i}\delta/2} \end{pmatrix}, \qquad (2)$$

where A_x and A_y are positive numbers that, through a suitable choice of units, can be assumed to satisfy the equality

$$A_x^2 + A_y^2 = 1, \qquad (3)$$

and where δ can be assumed as varying in the interval $0 \le \delta < 2\pi$.

It is noted that a symmetric form as in equation (2) is not necessary to represent polarization states, because in such a case one is interested only in the mutual dephasing between the two components, as already said. On the other hand, the symmetric choice may have some important implications in the problems of spinors, as will be clarified in the following.

Let us now observe that, thanks to equality (3), A_x and A_y can be thought of as the cosine and the sine (or vice versa) of an angular variable spanning the interval $[0, \pi/2]$. Let us tentatively set

$$A_x = \cos\gamma, \qquad A_y = \sin\gamma, \qquad (0 \le \gamma \le \pi/2). \qquad (4)$$

At this point it is rather natural to think that γ and δ could be taken as the colatitude and longitude, respectively, of a point on a unit sphere centred at the origin of a suitable reference frame, say ξ, η, ζ (of course such Cartesian coordinates should not be confused with the x, y, z coordinates in the space where the wave propagates). This leads to the representation of all possible states of polarization as points on the upper hemisphere. A few states are shown in figure 1. Such a representation, however, is slightly odd in that it does not treat x- and y-polarization on an equal footing. Indeed, while a linear polarization along the x-axis is represented by the north pole, a linear polarization along the y-axis corresponds to all the points on the equatorial line. Since in the latter case the phase difference δ between the x- and y-components of the electric field loses meaning, the only difference among points on the equatorial line would be the field initial phase, a quantity we decided to disregard. In order to eliminate this drawback we have to modify our representation in such a way that the whole equatorial line collapses into a single point. This is easily done on replacing equation (4) by

$$A_x = \cos\frac{\gamma}{2}, \qquad A_y = \sin\frac{\gamma}{2}, \qquad (0 \le \gamma \le \pi). \qquad (5)$$

The complete expression of a typical Jones vector now becomes

$$\begin{pmatrix} \cos(\gamma/2)\,\mathrm{e}^{-\mathrm{i}\delta/2} \\ \sin(\gamma/2)\,\mathrm{e}^{\mathrm{i}\delta/2} \end{pmatrix}, \qquad (0 \le \gamma \le \pi;\ 0 \le \delta < 2\pi). \qquad (6)$$

Again using γ as colatitude and δ as longitude, the set of points representing possible polarization states covers the entire surface of the unit sphere. In particular, linear y-polarization corresponds to the south pole. A few states are indicated in figure 2. A peculiar feature of this representation should be noted. States of orthogonal polarization are imaged onto points at opposite ends of a diametrical line. This can be seen in general, but we content ourselves with the self-evident case of linear polarization (we shall come back to this point in section 3). If we take the unit vector ending at a typical point of the sphere as representative of a polarization state, we can say that our mapping images orthogonal polarization states onto antiparallel unit vectors. Let us express the same result in a different form. Since any polarization state can be represented as a superposition of two orthogonal states, for example linear x- and y-polarization, any unit vector in the representation space can be thought of as the superposition of two antiparallel unit vectors along an arbitrarily chosen direction, for example the ζ-axis. At first, looking at the sphere in its three-dimensional ξ, η, ζ space, this seems surprising because we are accustomed to referring vectors to a frame of mutually orthogonal axes. On second thought, however, we realize that this is in fact the most profound meaning of our representation. Equation (6) can be read by saying that $\cos(\gamma/2)\,\mathrm{e}^{-\mathrm{i}\delta/2}$ and $\sin(\gamma/2)\,\mathrm{e}^{\mathrm{i}\delta/2}$ are the 'components' along the ζ-axis of the unit vector pointing in the direction (γ, δ).

As is well known, the Poincaré sphere is a classical way to represent polarization states [4, 5] and usually its introduction is made through the Stokes parameters, which are related to the normalized coordinates on the sphere. The sphere of figure 2 is similar to the Poincaré

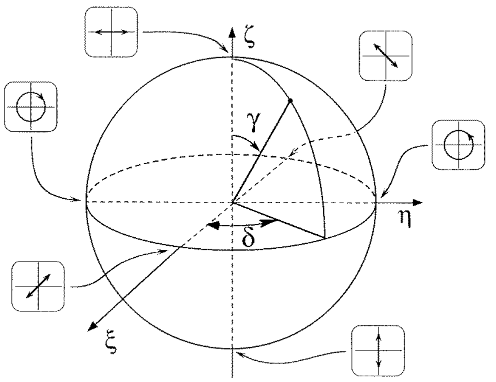

Figure 2. Light polarization states mapped on a sphere.

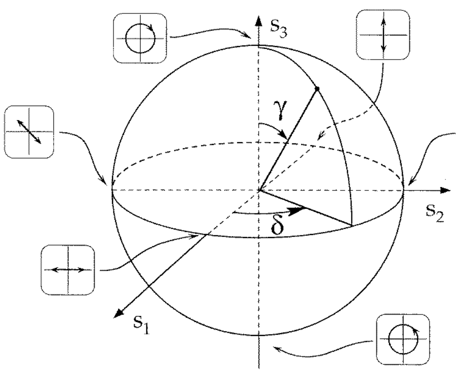

Figure 3. The Poincaré sphere.

sphere but it is not quite the same. For the sake of comparison, the Poincaré sphere is drawn in figure 3, where s_1, s_2 and s_3 are Stokes parameters [4]. It is seen, for example, that its poles represent states of circular instead of linear polarization. The reason for this difference is easily traced. The sphere of figure 2 originates from a representation by Jones vectors of the form (2) in which the basis vectors are

$$\begin{pmatrix}1\\0\end{pmatrix}, \qquad \begin{pmatrix}0\\1\end{pmatrix}, \tag{7}$$

corresponding to linearly-polarized fields along the x- and y-axis, respectively. On the other hand, any polarized wave can be represented as a suitable superposition of right- and left-circular-polarized fields. In this case the basis vectors are [4, 5]

$$\frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\pi/4}\\e^{i\pi/4}\end{pmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i3\pi/4}\\e^{i3\pi/4}\end{pmatrix}. \tag{8}$$

Starting from this new basis and proceeding as before we obtain the Poincaré sphere, except for a rotation by $\pi/2$ around the ζ-axis.

3. Spinors

The most important motivation for the introduction of spinors in physics is of course the existence of spin. To begin with, we recall a fundamental property of spin-$\frac{1}{2}$ particles. The measurement of spin along an arbitrary direction always leads to the results $+1$ or -1 (in $\hbar/2$ units). According to the principles of quantum mechanics this requires that any state of the particle be represented as a suitable superposition of the states leading to those results. In particular, if we know that the spin is along a direction having colatitude ϑ and longitude φ with respect to a given frame x, y, z, we should be able to represent it as a superposition of states in which the spin is in the positive or negative direction of the z-axis. The two latter states, which are often briefly called *spin up* and *spin down*, are orthogonal. Physically speaking, this means that they represent mutually exclusive outcomes. We are now confronted with the problem of how to represent spin states of the particle in mathematical terms. In particular, what type of representation leads to the orthogonality (in a mathematical sense) of *spin up* and *spin down*?

In view of the remarks made at the end of the previous section, we easily realize that column vectors of the form (6) are possible candidates. We shall know explore this possibility. In this exploration we will not use the spin and its physical properties.

Let us consider a typical unit vector specified in real space by colatitude ϑ and longitude φ. From now on we shall simply say the unit vector (ϑ, φ). We stipulate that such a unit vector be associated with the following mathematical object (of the same form as equation (6))

$$\begin{pmatrix}\cos(\vartheta/2)\,e^{-i\varphi/2}\\ \sin(\vartheta/2)\,e^{i\varphi/2}\end{pmatrix}, \qquad (0 \le \vartheta \le \pi\colon 0 \le \varphi < 2\pi). \tag{9}$$

Such an object will be called a *spinor*. One may well ask why we should introduce a new name instead of using the phrase *Jones vector* as was done in section 2. To answer this question we first observe that we are going to use the arguments of section 2 the other way around. There, we had, in physical space, a wave described in a rather obvious way by a Jones vector and we arrived at a *fictitious* three-dimensional space ξ, η, ζ where each Jones vector was associated to a unit vector (γ, δ). Here, we start with a physical three-dimensional space and we try to find a suitable representation of it through objects of the form (9). It could be objected that we are simply looking at the same matter from a different viewpoint. We then remark that such a different viewpoint makes relevant a question that would not appear the other way. How does a spinor change when we rotate the reference frame in real space? This is an important question. To understand why, let us recall that an ordinary (three-dimensional) vector is not just a triplet of real numbers but an abstract object represented by three numbers or, which is the same, by its three Cartesian components in a certain reference frame. The true nature of the vector is revealed by the transformation laws obeyed by those three numbers when the reference frame changes. The same is true for a spinor. As a mathematical object, a spinor is specified by its transformation rules under rotations of the reference frame. It is to find these rules that the next section is devoted. Before doing this, however, we need some more preparation. In particular, we must endow spinors with an inner product. If we denote two typical spinors by

$$\hat{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \qquad \hat{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{10}$$

where a_1, a_2 and b_1, b_2 are of the form (9), we define their inner (or dot) product as

$$\hat{a} \cdot \hat{b} = a_1^* b_1 + a_2^* b_2, \tag{11}$$

where the asterisk denotes complex conjugate. Note that interchanging the order of $\hat{a}$ and $\hat{b}$ leads to the complex conjugate. In particular, two spinors are said to be orthogonal if $\hat{a} \cdot \hat{b} = 0$.

In summary, our definitions are expressed by equation (9), which will be referred to as the association rule, and by equation (11), the inner product rule.

We can now control the fact that two opposite unit vectors on a line lead to orthogonal spinors. Let $\hat{a}$ be the spinor corresponding to the unit vector (ϑ, φ). More explicitly, we have

$$\begin{aligned} \hat{a} &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\vartheta/2)\, e^{-i\varphi/2} \\ \sin(\vartheta/2)\, e^{i\varphi/2} \end{pmatrix}. \end{aligned} \tag{12}$$

We denote by $\hat{a}_-$ the spinor obtained on passing to the opposite direction. This amounts to changing ϑ into $\pi - \vartheta$ and φ into $\varphi + \pi$. Using the association rule we find

$$\begin{aligned} \hat{a}_- &= \begin{pmatrix} \cos((\pi - \vartheta)/2)\, e^{-i(\varphi+\pi)/2} \\ \sin((\pi - \vartheta)/2)\, e^{i(\varphi+\pi)/2} \end{pmatrix} \\ &= \begin{pmatrix} -i a_2^* \\ i a_1^* \end{pmatrix}. \end{aligned} \tag{13}$$

On inserting equations (12) and (13) into the inner product rule (11) we see that $\hat{a} \cdot \hat{a}_- = 0$.

It will be useful to introduce a special notation for the spinors corresponding to unit vectors along the opposite directions on the Cartesian axes of real space. We denote such spinors by $\hat{x}, \hat{x}_-$, $\hat{y}, \hat{y}_-$ and $\hat{z}, \hat{z}_-$. On applying the association rule we easily find their expression:

$$\hat{x} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \hat{x}_- = \frac{-i}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{14}$$

$$\hat{y} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\pi/4} \\ e^{i\pi/4} \end{pmatrix}, \qquad \hat{y}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i3\pi/4} \\ e^{i3\pi/4} \end{pmatrix}, \tag{15}$$

$$\hat{z} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \hat{z}_- = \begin{pmatrix} 0 \\ i \end{pmatrix}. \tag{16}$$

It can be noted that equation (13) gives, for $\hat{z}_-$, a result which is not common in the literature, where the vector $\hat{z}_-$ is generally taken as (0, 1) [6–8]. This is due to the fact that the angle φ is indeterminate along the z-direction (we have arbitrarily set φ equal to zero). It could be easily verified that the use of either of the two spinors would lead to identical results as far as their transformation properties is concerned. However, for the sake of mathematical consistency, we shall use the definition of $\hat{z}_-$ given in equation (16).

Let us remark that a typical spinor can be seen as a combination of $\hat{z}$ and $\hat{z}_-$. In fact we can write

$$\hat{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} - i a_2 \begin{pmatrix} 0 \\ i \end{pmatrix} = a_1 \hat{z} - i a_2 \hat{z}_-. \tag{17}$$

This is an important formula. It shows how the spinor associated with an arbitrary direction is expressed through the spinors of the z-axis. Remember that representing any direction in space by means of opposite directions on one and the same line was one of our starting requirements. The complex numbers a_1 and $-ia_2$ are the components of $\hat{a}$ on $\hat{z}$ and $\hat{z}_-$, respectively. On the other hand, $\hat{a}$ can also be expressed through $\hat{x}$, $\hat{x}_-$ or $\hat{y}$, $\hat{y}_-$. To see how, note that, taking equation (14) into account, $\hat{x}$ and $\hat{x}_-$ can be written

$$\begin{aligned} \hat{x} &= \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} - i \begin{pmatrix} 0 \\ i \end{pmatrix} \right] = \frac{1}{\sqrt{2}} (\hat{z} - i\hat{z}_-), \\ \hat{x}_- &= \frac{-i}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ i \end{pmatrix} \right] = \frac{-i}{\sqrt{2}} (\hat{z} + i\hat{z}_-). \end{aligned} \tag{18}$$

These two equations are easily inverted to give

$$\hat{z} = \frac{1}{\sqrt{2}} (\hat{x} + i\hat{x}_-), \qquad \hat{z}_- = \frac{i}{\sqrt{2}} (\hat{x} - i\hat{x}_-). \tag{19}$$

On inserting equation (19) into equation (17) we obtain

$$\hat{a} = \frac{1}{\sqrt{2}} [(a_1 + a_2)\hat{x} + i(a_1 - a_2)\hat{x}_-]. \tag{20}$$

This formula shows that the components of $\hat{a}$ on $\hat{x}$ and $\hat{x}_-$ are $(a_1 + a_2)/\sqrt{2}$ and $i(a_1 - a_2)/\sqrt{2}$, respectively. With an analogous procedure we could obtain the expressions for the components on $\hat{y}$ and $\hat{y}_-$.

4. Spinors under rotations

We begin our study about the effect on spinors of rotating the reference frame by starting from the simplest case: a rotation about the z-axis by an angle α. Let us take an arbitrary unit vector (ϑ, φ) in the real space. Then a certain spinor $\hat{a}$ is associated with it. In figure 4 we show by a dashed line the intersection between the x, y plane and the plane passing through the z-axis and containing the chosen unit vector. It is seen that in the rotated frame x', y' the longitude of the unit vector becomes

$$\varphi' = \varphi - \alpha. \tag{21}$$

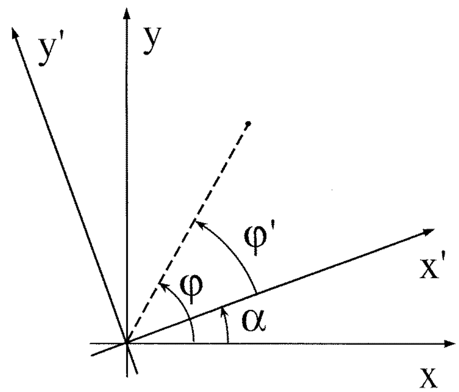

Figure 4. Notation used for the analysis of the rotation around the z-axis.

According to the association rule (9) the corresponding spinor, say $\hat{a}'$, has the explicit form

$$\hat{a}' = \begin{pmatrix} \cos(\vartheta/2)\, e^{-i(\varphi-\alpha)/2} \\ \sin(\vartheta/2)\, e^{i(\varphi-\alpha)/2} \end{pmatrix}. \tag{22}$$

This result can be synthesized in matrix form by writing

$$\begin{pmatrix} a'_1 \\ a'_2 \end{pmatrix} = R_z(\alpha) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \tag{23}$$

where

$$R_z(\alpha) = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} \tag{24}$$

is called the rotation matrix for the z-axis.

We now ask how this result is to be extended to rotations about other axes. As before, let us refer for example to the x-axis. Is the rotation matrix still the same? The answer is negative because the z-axis is somewhat privileged. We know in fact that in the expression of $\hat{a}$ the numbers a_1 and $-ia_2$ are the components with respect to $\hat{z}$ and $\hat{z}_-$ (see equation (17)). On the other hand, we know how to express $\hat{a}$ through $\hat{x}$ and $\hat{x}_-$ by means of equation (20). Therefore, we can say that the rotation matrix for the x-axis has the form appearing in equation (24) if we apply it to the components of $\hat{a}$ with respect to $\hat{x}$ and $\hat{x}_-$ (instead of $\hat{z}$ and $\hat{z}_-$). As a consequence, we can write

$$\frac{a'_1 + a'_2}{\sqrt{2}} = \frac{a_1 + a_2}{\sqrt{2}}\, e^{i\alpha/2}, \tag{25}$$

$$i\frac{a'_1 - a'_2}{\sqrt{2}} = i\frac{a_1 - a_2}{\sqrt{2}}\, e^{-i\alpha/2}. \tag{26}$$

Solving with respect to a'_1 and a'_2 we find

$$a'_1 = a_1 \cos\frac{\alpha}{2} + ia_2 \sin\frac{\alpha}{2}, \tag{27}$$

$$a'_2 = ia_1 \sin\frac{\alpha}{2} + a_2 \cos\frac{\alpha}{2}. \tag{28}$$

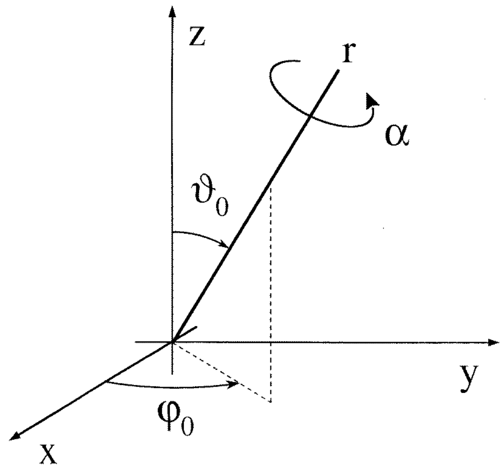

Figure 5. Notation used for the analysis of the rotation around an arbitrary axis r.

The link between the components of $\hat{a}$ before and after the rotation can be expressed in matrix form by

$$\begin{pmatrix} a'_1 \\ a'_2 \end{pmatrix} = R_x(\alpha) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \tag{29}$$

where the rotation matrix with respect to x is

$$R_x(\alpha) = \begin{pmatrix} \cos(\alpha/2) & i\sin(\alpha/2) \\ i\sin(\alpha/2) & \cos(\alpha/2) \end{pmatrix}. \tag{30}$$

We are now ready to generalize the previous results to rotation by an angle α about an arbitrary axis, say r, specified by colatitude ϑ_0 and longitude φ_0 as sketched in figure 5.

Denote by $\hat{r}$ and $\hat{r}_-$ the spinors associated with r. Taking into account equations (17) and (13) they can be written as

$$\hat{r} = r_1\hat{z} - ir_2\hat{z}_-, \qquad \hat{r}_- = -ir_2^*\hat{z} + r_1^*\hat{z}_-, \tag{31}$$

where, according to the association rule, we have

$$r_1 = \cos(\vartheta_0/2)\, e^{-i\varphi_0/2}, \qquad r_2 = \sin(\vartheta_0/2)\, e^{i\varphi_0/2}. \tag{32}$$

Solving equation (31) with respect to $\hat{z}, \hat{z}_-$ we obtain

$$\hat{z} = r_1^*\hat{r} + ir_2\hat{r}_-, \qquad \hat{z}_- = ir_2^*\hat{r} + r_1\hat{r}_-. \tag{33}$$

Let us now insert equation (33) into equation (17). The expression for a typical spinor $\hat{a}$ then becomes

$$\hat{a} = (a_1r_1^* + a_2r_2^*)\hat{r} + i(a_1r_2 - a_2r_1)\hat{r}_-. \tag{34}$$

Equation (34) expresses $\hat{a}$ as a combination of $\hat{r}$ and $\hat{r}_-$. Then the line r has taken on the role of the z-axis. The components of $\hat{a}$ with respect to $\hat{r}$ and $\hat{r}_-$ are $(a_1r_1^* + a_2r_2^*)$ and $i(a_1r_2 - a_2r_1)$, respectively. Upon rotation of α about r they change through multiplication by the matrix (24) or

$$(a'_1r_1^* + a'_2r_2^*) = (a_1r_1^* + a_2r_2^*)\, e^{i\alpha/2}, \tag{35}$$

$$i(a'_1r_2 - a'_2r_1) = i(a_1r_2 - a_2r_1)\, e^{-i\alpha/2}. \tag{36}$$

These equations can now be solved with respect to a_1' and a_2'. The result is

$$a_1' = a_1[|r_1|^2 e^{i\alpha/2} + |r_2|^2 e^{-i\alpha/2}] + 2ia_2 r_1 r_2^* \sin(\alpha/2), \tag{37}$$

$$a_2' = 2ia_1 r_1^* r_2 \sin(\alpha/2) + a_2[|r_1|^2 e^{-i\alpha/2} + |r_2|^2 e^{i\alpha/2}]. \tag{38}$$

The resulting rotation matrix is

$$R_r(\alpha) = \begin{pmatrix} |r_1|^2 e^{i\alpha/2} + |r_2|^2 e^{-i\alpha/2} & 2ir_1 r_2^* \sin(\alpha/2) \\ 2ir_1^* r_2 \sin(\alpha/2) & |r_1|^2 e^{-i\alpha/2} + |r_2|^2 e^{i\alpha/2} \end{pmatrix}. \tag{39}$$

Taking equation (32) into account and using trigonometric identities, equation (39) can be written

$$R_r(\alpha) = \begin{pmatrix} \cos(\alpha/2) + i\cos\vartheta_0 \sin(\alpha/2) & (i\sin\vartheta_0\cos\varphi_0 + \sin\vartheta_0\sin\varphi_0)\sin(\alpha/2) \\ (i\sin\vartheta_0\cos\varphi_0 - \sin\vartheta_0\sin\varphi_0)\sin(\alpha/2) & \cos(\alpha/2) - i\cos\vartheta_0\sin(\alpha/2) \end{pmatrix}. \tag{40}$$

If the direction cosines of r,

$$r_x = \sin\vartheta_0\cos\varphi_0, \qquad r_y = \sin\vartheta_0\sin\varphi_0, \qquad r_z = \cos\varphi_0, \tag{41}$$

are used, equation (40) becomes

$$R_r(\alpha) = \begin{pmatrix} \cos(\alpha/2) + ir_z\sin(\alpha/2) & (ir_x + r_y)\sin(\alpha/2) \\ (ir_x - r_y)\sin(\alpha/2) & \cos(\alpha/2) - ir_z\sin(\alpha/2) \end{pmatrix}. \tag{42}$$

This is the matrix describing the most general rotation. A spinor can now be characterized as an object of the form (12) (in a certain frame) that under rotation of the coordinate axes changes through multiplication by $R_r(\alpha)$.

There is a significant consequence of the rotation law that is worth noting. If the angle α equals 2π the rotation matrix becomes the opposite of the identity matrix. This means that under such a rotation both components of any spinor change sign. Since a 2π-rotation brings back the reference frame to its initial configuration, such a result could be considered as an inconsistency of our association rule. The surprise comes when spinors are applied to particles with spin. It turns out that upon 2π-rotation of the particle spin (which is tantamount to rotating the reference frame in the opposite sense), the change of sign for the components of the corresponding spinor should lead to detectable effects in certain physical situations. These predictions have been confirmed through suitable experiments [9]. It is interesting to remark that the change of sign is a consequence of the choice leading to the symmetric form in equation (2). This choice is arbitrary from a purely mathematical point of view, but seems to be the correct one to take into account this experimentally evidenced phenomenon. On the other hand, it should be noted that the topic concerning the change of sign of a spinor due to rotations by 2π is still under discussion [10].

5. Pauli matrices

The expression obtained in the previous section for the most general rotation can be given a different and useful form. The matrix appearing in equation (42) can in fact be written

$$R_r(\alpha) = \cos(\alpha/2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i\sin(\alpha/2)\left[r_x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + r_y\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + r_z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right]. \tag{43}$$

Of the four 2×2 matrices appearing on the right, the first one is of course the identity matrix, to be denoted by I, while the other three, denoted by σ_x, σ_y and σ_z, are the celebrated Pauli matrices [6]. More explicitly, we have

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{44}$$

As could be simply verified, the Pauli matrices obey the following rules,

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l, \qquad \{\sigma_j, \sigma_k\} = 2\delta_{jk}I, \qquad (j, k, l = x, y, z), \tag{45}$$

where square and curly brackets denote commutator and anticommutator, respectively. ϵ_{jkl} is the Levi-Civita antisymmetric tensor and δ_{jk} is the Kronecker symbol.

With the present notation, equation (43) can be synthesized as follows,

$$R_r(\alpha) = \cos(\alpha/2)I + i\sin(\alpha/2)[r_x\sigma_x + r_y\sigma_y + r_z\sigma_z] = \cos(\alpha/2)I + i\sin(\alpha/2)\boldsymbol{r}\cdot\boldsymbol{\sigma}, \tag{46}$$

where $\boldsymbol{r}$ is the unit vector, with components r_x, r_y and r_z along the line r, and $\boldsymbol{\sigma}$ stands for a sort of matrix vector whose components are the Pauli matrices. Let us transform equation (46) into a symbolic form that turns out to be useful in quantum mechanics. To this end we write down the well known series expansions

$$\cos(\alpha/2) = \sum_{k=0}^{\infty}\frac{(i\alpha/2)^{2k}}{(2k)!}, \qquad i\sin(\alpha/2) = \sum_{k=0}^{\infty}\frac{(i\alpha/2)^{2k+1}}{(2k+1)!}. \tag{47}$$

On inserting from the latter equation into equation (46) we obtain

$$R_r(\alpha) = \sum_{k=0}^{\infty}\left[\frac{(\alpha/2)^{2k}(i)^{2k}I}{(2k)!} + \frac{(\alpha/2)^{2k+1}(i)^{2k+1}\boldsymbol{r}\cdot\boldsymbol{\sigma}}{(2k+1)!}\right]. \tag{48}$$

It is easily seen by using ordinary matrix multiplication that

$$(\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{\sigma})^{2k} = (\mathrm{i})^{2k} I, \qquad (\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{\sigma})^{2k+1} = (\mathrm{i})^{2k+1}\boldsymbol{r}\cdot\boldsymbol{\sigma}. \tag{49}$$

When such identities are inserted into equation (48) the following expression is obtained:

$$R_r(\alpha) = \sum_{k=0}^{\infty} \frac{(\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{\sigma}\alpha/2)^k}{k!} = \exp\left(\frac{\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{\sigma}\alpha}{2}\right). \tag{50}$$

The expression on the right-hand side of equation (50) is just the same as that which is frequently used in quantum mechanics to describe rotation of the spin. Actually it often represents the starting point to discuss the effects of rotation, being in many cases derived from the infinitesimal rotation operator.

References

[1] Cartan E 1966 *The Theory of Spinors* (Cambridge, MA: MIT)
[2] Frescura F A M and Hiley B J 1981 *Am. J. Phys.* **49** 152
[3] Payne W T 1952 *Am. J. Phys.* **20** 253
[4] Born M and Wolf E 1991 *Principles of Optics* 6th edn (Oxford: Pergamon)
[5] Shurcliff W A 1962 *Polarized Light* (Cambridge, MA: Harvard University Press)
[6] Sakurai J J 1985 *Modern Quantum Mechanics* (Reading, MA: Addison-Wesley)
[7] Merzbacher E 1970 *Quantum Mechanics* 2nd edn (New York: Wiley)
[8] Landau L D and Lifshitz E M 1958 *Quantum Mechanics: Non relativistic Theory* (London: Pergamon)
[9] Rauch H 1986 *Contemp. Phys.* **27** 345
[10] Gough W 1992 *Eur. J. Phys.* **13** 167