Pramāṇa – J. Phys., Vol. 32, No. 6, June 1989, pp. 769–792. © Printed in India.

Hamilton's theory of turns and a new geometrical representation for polarization optics

R SIMON*, N MUKUNDA† and E C G SUDARSHAN§

Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560 012, India
*Present address: The Institute of Mathematical Sciences, C.I.T. Campus, Madras 600 113, India
§Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA

MS received 1 March 1989.

Abstract. Hamilton's theory of turns for the group SU(2) is exploited to develop a new geometrical representation for polarization optics. While pure polarization states are represented by points on the Poincaré sphere, linear intensity preserving optical systems are represented by great circle arcs on another sphere. Composition of systems, and their action on polarization states, are both reduced to geometrical operations. Several synthesis problems, especially in relation to the Pancharatnam-Berry-Aharonov-Anandan geometrical phase, are clarified with the new representation. The general relation between the geometrical phase, and the solid angle on the Poincaré sphere, is established.

Keywords. Polarization optics; geometrical phases; theory of turns; Poincaré sphere; Pancharatnam phase.

PACS Nos 42·10; 02·20; 42·78

1. Introduction

The recognition by Poincaré that the polarization states of a plane electromagnetic wave can be made to correspond one-to-one to the points on a two-dimensional sphere S^2 represents a physical and mathematical fact of deep significance (Poincaré 1892; Ramachandran and Ramaseshan 1961; Born and Wolf 1965). It allows one to represent different polarization states which are close to each other in an obvious physical sense, by points on S^2 which are near to each other in the obvious topological sense. With increased sophistication, one now understands the relevance of S^2 in this context by realising that it is the coset space $SU(2)/U(1) = S^3/S^1$.

The importance of the group SU(2) in polarization optics stems basically from the fact that it acts in a natural way on its coset space $SU(2)/U(1) \simeq S^2$. (Its relevance in a variety of quantum mechanical problems needs no emphasis). Stated in physical terms, problems involving polarization states of plane light waves and their transformation by intensity preserving linear optical systems lead one immediately to SU(2). These problems have received much attention recently, particularly in connection with the Berry, Aharonov-Anandan phase (Berry 1984, 1985, 1987a, b; Aharonov and Anandan 1987) and its relationship to a phase discovered by Pancharatnam in the fifties (Pancharatnam 1956).

†Jawaharlal Nehru Fellow

Hamilton, in the course of his studies on quaternions, discovered in 1853 a geometrical or pictorial way of representing the elements of the group SU(2) and their (noncommutative) composition law (Hamilton 1853). To appreciate Hamilton's method, it is appropriate to recall first the elementary pictorial representation of the much simpler abelian group of translations in Euclidean space. Elements of this group can be pictured as free vectors which are directed line segments in Euclidean space. A free vector is an equivalence class of ordinary vectors with location, two ordinary vectors being equivalent if by a rigid translation one can be made to coincide with the other. Then a faithful graphical description of the composition law for the translation group is given by the parallelogram law for addition of free vectors: given two vectors, move the second rigidly, i.e. parallel to itself, till its tail coincides with the head of the first vector; then the resultant vector runs from the tail of the first to the head of the second.

Hamilton's work generalises this pictorial representation from the abelian translation group to the nonabelian group SU(2), and is based on the concept of turns. A turn is an equivalence class of directed great circle arcs on the unit sphere S^2: the arcs obtained by sliding a given arc over its great circle by all possible amounts are all equivalent to one another. One can now associate each SU(2) element with a turn in a natural way; and the composition law for SU(2) can be correctly reproduced by a noncommutative geometric "addition" or composition rule for turns, analogous to the parallelogram law for translation vectors. Thus, to obtain the product of two SU(2) elements, one slides the corresponding turns on their respective great circles (more precisely, one chooses representative arcs) to one of the points where the two circles intersect, in such a way that the tail of the second turn coincides with the head of the first. Then the turn determined by the great circle arc from the tail of the first turn to the head of the second corresponds to the product of the two group elements. A lucid and comprehensive account of this remarkably elegant geometrical description of SU(2) has been given by Biedenharn and Louck (1981).

The purpose of the present work is to bring out the role played by the group SU(2) in polarization optics, and to employ Hamilton's theory of turns to develop a new geometrical representation for this important branch of optics.

The plan of the paper is as follows. In §2 we review some basic concepts from polarization optics with the aim of highlighting the role of SU(2) in this context. The Poincaré sphere method is seen to be a physical realization of the two-to-one homomorphism from SU(2) to SO(3). Given a general Jones matrix $J \in SU(2)$, the SO(3) rotation Ω (J) it produces on the Poincaré sphere is presented in a simple form which we believe is new in the context of polarization optics.

In §3 we reformulate Hamilton's method of turns in a form convenient for the present purpose, namely with the aid of algebraic expressions which make the geometrical properties evident, and derive the composition rule for turns in a straightforward manner. This immediately leads to our new geometrical representation for polarization optics based on two unit spheres S^2, the Poincaré sphere $\mathscr{P}$ and the sphere of turns $\mathscr{T}$. States and their transformations are geometrically represented on $\mathscr{P}$, while linear intensity preserving optical systems are represented as turns on $\mathscr{T}$. The composition of the action of a sequence of optical systems is then handled geometrically rather than algebraically.

Some applications of the new representation are presented in §4. The synthesis of optical rotators using a pair of half wave plates is analyzed, and the importance of such a synthesis in the experimental measurement of the Aharonov-Anandan geometrical phase is brought out. We also describe the synthesis of a gadget involving only birefringent plates which is able to realize all SU(2) transformations through simple rotations of the components. The suggestive power of the method of turns will be quite evident from these examples of system synthesis.

In §5 we derive the relationship between the geometric phase and the solid angle subtended by a closed circuit on the Poincaré sphere $\mathscr{P}$. An example where this relationship is dynamically manifested is discussed. Finally, in §6 we present some concluding remarks.

2. Polarization states, Poincaré Sphere and SU(2)

The polarization state of a fully polarized quasimonochromatic plane wave propagating along the x_3 direction is given by the Maxwell vector (also called Jones vector)

$$E = \begin{pmatrix} E_1 \\ E_2 \end{pmatrix}, \tag{1}$$

where the complex components E_1, E_2 of the transverse electric vector in the directions x_1, x_2 are independent of the transverse location x_1, x_2. An optical system which transforms the polarization state linearly is described by a 2×2 matrix J with complex entries:

$$E \rightarrow E' = JE \tag{2}$$

J is called the Jones matrix of the system. If the transformation preserves the intensity $I = E^{+}E$, it follows that J must be an element of the unitary group U(2). Thus, there is a one-to-one correspondence between intensity preserving linear systems and elements of U(2): the action of two systems in succession is given by the product of the associated U(2) matrices in that order.

If we were not interested in the overall phase, then each optical system could be effectively represented by an SO(3) rotation on the Poincaré sphere $\mathscr{P}$. However, we *are* interested in the overall phase, in addition to other things; this phase plays an important role in, for example, the Berry-Pancharatnam-Aharonov-Anandan kinds of situations which have received much attention recently.

Keeping the above in mind, we represent each U(2) optical system J by an ordered pair (φ, u), where φ is a $U(1)$ phase angle and $u \in \mathrm{SU}(2)$. Since $\mathrm{U}(2) = ((\mathrm{U}(1) \times \mathrm{SU}(2))/Z_2$, we see that (φ, u) and $(\varphi + \pi, -u)$ represent one and the same U(2) optical system, so there is a two-fold ambiguity in representing J by a pair (φ, u). Thus in the composition of m optical systems $J_1, J_2, \ldots, J_m \in \mathrm{U}(2)$, there are 2^m distinct ways of representing this as the composition of m pairs; each one of them of course leads to the same final U(2) element for the product system $J = J_m \cdots J_2 J_1$. If J corresponds to the pair $(\varphi, u) \equiv (\varphi + \pi, -u)$, for each choice of pairs $(\varphi_1, u_1), \ldots, (\varphi_m, u_m)$, the total phase $\varphi_1 + \varphi_2 + \cdots + \varphi_m$ equals either φ or $\varphi + \pi$, and correspondingly the SU(2) product $u_m \cdots u_2 u_1$ equals u or $-u$. With such decompositions, and the above equivalence understood, we are free to use Hamilton's method of turns in calculating the composition of SU(2) elements, with the abelian phases kept track of appropriately. For brevity and simplicity of notation, we will occasionally represent the (two-valued) SU(2) part of a U(2) optical system itself by J. In view of the above discussion this should cause no confusion.

Typical examples of linear systems of interest in polarization optics are the rotator, and the compensator or birefringent plate. The rotator has a Jones matrix

$$J=\exp(i\varphi)R_\theta=\exp(i\varphi)\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}, \tag{3}$$

and has the effect of rotating the Cartesian components E_1, E_2 through an angle θ. The compensator, characterised by mutually orthogonal slow and fast axes, advances the phase of the component of E along the former by an amount δ relative to the component along the latter. If the fast axis is along x_1, the Jones matrix is

$$J=\exp(i\chi)(C_\delta)_0=\exp(i\chi)\begin{pmatrix}e^{-i\delta/2} & 0\\ 0 & e^{i\delta/2}\end{pmatrix}. \tag{4}$$

For $\delta=\pi/2$ or π we have a quarterwave plate or halfwave plate respectively. For a general position of the fast axis, at an angle θ with the x_1 axis, we have

$$\begin{aligned}J&=\exp(i\chi)(C_\delta)_\theta=\exp(i\chi)R_\theta(C_\delta)_0R_\theta^{-1}\\ &=\exp(i\chi)\begin{pmatrix}\cos\delta/2-i\cos 2\theta\sin\delta/2 & -i\sin 2\theta\sin\delta/2\\ -i\sin 2\theta\sin\delta/2 & \cos\delta/2+i\cos 2\theta\sin\delta/2\end{pmatrix}.\end{aligned} \tag{5}$$

Here we have chosen $\exp(i\varphi)$ and $\exp(i\chi)$ as the $U(1)$ parts of these systems. Free propagation, and passage through a nonbirefringent and optically inactive medium, are examples of pure $U(1)$ systems (with the SU(2) parts taken to be the identity), as they simply increase the phases of E_1 and E_2 by equal amounts. One way to realize the rotator is by propagation through an optically active medium of appropriate specific activity and thickness, with φ and θ both being proportional to the thickness. Another often convenient way, as we shall see in the sequel, is by using a pair of (birefringent) half wave plates.

In analogy with the pure state density matrix in quantum mechanics, one defines the coherency or polarization matrix Φ corresponding to a Maxwell vector E as

$$\Phi=EE^\dagger. \tag{6}$$

Since we are concerned with unitary transformations which preserve the total intensity, we can conveniently normalize E to have unit intensity. This makes Φ a 2×2 projection matrix:

$$\Phi^\dagger=\Phi,\quad \operatorname{tr}\Phi=1,\quad \det\Phi=0. \tag{7}$$

By expanding Φ in terms of the Pauli matrices, we make contact with the Poincaré sphere representation. To be consistent with the usual conventions in optics, where the circular polarization states lie at the poles of the Poincaré sphere, we adopt the following representation of the Pauli algebra:

$$\rho_1=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\quad \rho_2=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \rho_3=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}. \tag{8}$$

This is related to the usual quantum mechanical choice by a cyclic permutation, so we have

$$\rho_j\rho_k=\delta_{jk}1+i\varepsilon_{jkl}\rho_l. \tag{9}$$

The expansion of Φ in terms of $\boldsymbol{\rho}$ brings in a vector $\hat{n}$:

$$\Phi = \tfrac{1}{2}(\mathbb{1} + \hat{n}\cdot\boldsymbol{\rho}),$$

$$n_j = \operatorname{tr}\Phi\rho_j. \tag{10}$$

The hermiticity of Φ makes $\hat{n}$ real, and the condition $\det\Phi = 0$ makes $\hat{n}$ a unit vector. Thus pure polarization states are in one-to-one correspondence with points $\hat{n}$ on the unit sphere S^2 in Euclidean three-space, and this is the Poincaré sphere $\mathscr{P}$. For pure states, points on $\mathscr{P}$ are also called Stokes vectors (Stokes 1852). Circular polarization states correspond to the poles and linear polarizations to points on the equator. Other points on $\mathscr{P}$ represent general elliptic polarizations. For two general (normalized) Maxwell vectors E, E', the corresponding coherency matrices and points on $\mathscr{P}$ obey

$$\operatorname{tr}\Phi'\Phi = |E'^{\dagger}E|^2 = \tfrac{1}{2}(1 + \hat{n}', \hat{n}). \tag{11}$$

Thus diametrically opposite points on $\mathscr{P}$ represent mutually orthogonal polarization states. In all this, the conserved intensity and overall phase of E are suppressed.

A general state of elliptic polarization with azimuth θ, $0 \leqslant \theta \leqslant \pi$ (this is the angle between the major axis and the x_1 direction) and ellipticity angle ε, $-\pi/4 \leqslant \varepsilon \leqslant \pi/4$ ($\tan\varepsilon$ = ellipticity = ratio of minor to major axis, handedness = sign of ε) has Maxwell vector

$$E = A\exp(i\delta)\begin{pmatrix}\cos\theta\cos\varepsilon - i\sin\theta\sin\varepsilon\\ \sin\theta\cos\varepsilon + i\cos\theta\sin\varepsilon\end{pmatrix}, \tag{12}$$

where A and δ are real. The corresponding point on $\mathscr{P}$—the Stokes vector—is independent of A and δ:

$$\hat{n} = (\cos 2\theta\cos 2\varepsilon, \sin 2\theta\cos 2\varepsilon, \sin 2\varepsilon). \tag{13}$$

In particular, for right circular polarization, $\hat{n}$ is the north pole on $\mathscr{P}$:

$$E = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix} \Rightarrow \hat{n} = (0, 0, 1). \tag{14}$$

When E changes according to eq. (2), Φ transforms as

$$\Phi \to \Phi' = J\Phi J^{\dagger}, \tag{15}$$

where $J \in U(2)$. Then, as is well known, $\hat{n} \in \mathscr{P}$ undergoes a rotation $\Omega\,(J) \in SO(3)$. As explained earlier in this section, we can focus our attention on the (two-valued) SU(2) part of the system, which for brevity we have agreed to denote again by J. Then J can be expressed as an exponential

$$J = \exp\left(\frac{-i}{2}\alpha(J)\hat{\alpha}(J)\cdot\boldsymbol{\rho}\right), \tag{16}$$

with $\alpha(J)$ a real scalar and $\hat{\alpha}(J)$ a unit vector. From the properties (9) of the Pauli algebra, one finds

$$J\hat{n}\cdot\boldsymbol{\rho}J^\dagger = \hat{n}'\cdot\boldsymbol{\rho},$$

$$n'_j = \Omega_{jk}(J)n_k,$$

$$\Omega_{jk}(J) = \delta_{jk}\cos\alpha(J) + \hat{a}_j(J)\hat{a}_k(J)(1-\cos\alpha(J)) - \varepsilon_{jkl}\hat{a}_l(J)\sin\alpha(J). \tag{17}$$

Thus $\Omega(J)$ rotates $\hat{n}$ through an angle $\alpha(J)$ about the axis $\hat{a}(J)$, and the unitary action of an optical system on E results in an orthogonal rotation on $\mathscr{P}$. The fact that every unitary Jones matrix has two orthogonal eigenpolarization states reduces to Euler's theorem on rotations: the eigenstates correspond to the diametrically opposite points at which the axis of rotation intersects $\mathscr{P}$, namely $\hat{n} = \pm\hat{a}(J)$. Suppressing the $U(1)$ factor, we see for example that for the optical rotator (3) we have $\hat{a} = (0,0,1)$ and $\alpha = 2\theta$: thus on $\mathscr{P}$ it produces a rotation of amount 2θ about the positive n_3 axis, and has the states of circular polarization as eigenstates. Similarly, the compensator (5) has $\hat{a} = (\cos 2\theta, \sin 2\theta, 0)$, $\alpha = \delta$. It rotates $\mathscr{P}$ through an angle δ about the axis in the equatorial plane at an angle 2θ to the n_1 axis. Its eigenstates are represented by the two diametrically opposite points in the equatorial plane of $\mathscr{P}$ at angles 2θ, $2\theta+\pi$; in physical space these correspond to linear polarizations at angles θ, $\theta+\pi/2$ to the x_1 axis.

In the general case, the exponential form (16) for J is related to an expansion of J itself in terms of the Pauli matrices in this way:

$$J = a_0 - i\mathbf{a}\cdot\boldsymbol{\rho} = \begin{pmatrix} a_0 - ia_1 & -a_3 - ia_2 \\ a_3 - ia_2 & a_0 + ia_1 \end{pmatrix}$$

$$= \exp\left(\frac{-i}{2}\alpha\hat{a}\cdot\boldsymbol{\rho}\right) = \cos\frac{\alpha}{2} - i\hat{a}\cdot\boldsymbol{\rho}\sin\frac{\alpha}{2},$$

$$a_0 = \cos\frac{\alpha}{2}, \quad \mathbf{a} = \hat{a}\sin\frac{\alpha}{2}, \quad a_0^2 + \mathbf{a}\cdot\mathbf{a} = 1. \tag{18}$$

Thus from a_0 and $\mathbf{a}$ we can read off the values of α and $\hat{a}$. Noting that ($\hat{a}$, $\alpha/2$) and $(-\hat{a}, 2\pi-\alpha/2)$ lead to the same J, we can obtain all elements of SU(2) by restricting $\alpha/2$ to the range $0 \leqslant \alpha/2 \leqslant \pi$, but allowing $\hat{a}$ to assume all values on S^2. So we have:

$$\frac{\alpha}{2} = \cos^{-1} a_0, \quad 0 \leqslant \frac{\alpha}{2} \leqslant \pi,$$

$$\hat{a} = \mathbf{a}/\sqrt{1-a_0^2}, \tag{19}$$

determining the axis $\hat{a}$ about which J produces a rotation on $\mathscr{P}$.

To conclude this section we note that the connection between the unitary action of optical systems on the Maxwell vector E, and the transformations they produce on $\mathscr{P}$, is a concrete optical illustration of the two-to-one homomorphism from SU(2) to SO(3), which is also important in quantum mechanics. Here it is seen in the fact that the Jones matrices J and $-J$ produce the same rotation on $\mathscr{P}$. Rotations about the n_3 axis correspond to optically active media, and those about axes on the equatorial plane to birefringent plates with various positions of fast (and slow) axes. Rotations about other directions correspond to passage of light through media having both birefringence and optical activity: for a generator proportional to $\hat{a}\cdot\boldsymbol{\rho}$, the ratio between the specific optical activity and the specific birefringence is $\hat{a}_3/(\hat{a}_1^2+\hat{a}_2^2)^{1/2}$.

3. Hamilton's turns and the new representation for polarization optics

In §2 we have reviewed the fundamental role of the group SU(2) in polarization optics involving intensity preserving systems. We have also mentioned Hamilton's elegant method of turns in giving a geometrical representation for SU(2) and its composition law. Putting the two together we are led to use turns to obtain a new geometrical representation for polarization optics. We develop such a representation in this section, and begin with a formulation of Hamilton's ideas in a convenient form. A detailed exposition can be found in the work of Biedenharn and Louck (1981) referred to earlier; our approach is slightly different, and in particular readily lends itself to generalization to the non compact group SU(1, 1) as shown elsewhere. (Simon *et al* 1988).

The description of a general $J\in$SU(2) with the help of the homogeneous Euler parameters a_0, $\mathbf{a}$ in eq. (18) shows that there is a one-to-one correspondence between elements of SU(2) and points on S^3, the unit sphere in Euclidean four-space. However the expressions for a_0 and $\mathbf{a}$ in terms of α and $\hat{a}$ naturally suggest a way of associating elements of SU(2) with ordered pairs of points of S^2, the unit sphere in Euclidean three-space; it is this that leads to turns.

Let $\hat{n}$, $\hat{n}'$ be unit vectors on S^2 and let us set

$$a_0=\hat{n}\cdot\hat{n}',\quad \mathbf{a}=\hat{n}_\wedge\hat{n}'. \tag{20}$$

We do obtain in this way a point on S^3, and so we can associate a definite SU(2) element $A(\hat{n},\hat{n}')$ with the ordered pair $(\hat{n},\hat{n}')$ on S^2:

$$A(\hat{n},\hat{n}')=\hat{n}\cdot\hat{n}'-i\hat{n}_\wedge\hat{n}'\cdot\boldsymbol{\rho}. \tag{21}$$

Equivalently we can associate $A(\hat{n},\hat{n}')$ with the directed great circle arc from $\hat{n}$ to $\hat{n}'$ on S^2 (the arc length is always assumed to be $\leqslant\pi$). It should however be obvious that this association is not unique, since any other pair of unit vectors on the same great circle, obtained by rotating $\hat{n}$ and $\hat{n}'$ through the same angle about $\hat{n}_\wedge\hat{n}'$, leads to the same SU(2) element. Such a rotation corresponds to sliding the directed arc on its great circle.

This construction leads to Hamilton's turns. A turn is an equivalence class of directed great circle arcs: two arcs are equivalent if one is obtained from the other by sliding it along the great circle. Thus (except for $J=-1$) with every $J\in$SU(2) we can associate a unique turn. With J given in terms of homogeneous Euler parameters, $\mathbf{a}$ specifies the great circle as the "equator" with respect to it; the sense of the turn is also determined by $\mathbf{a}$ via the right hand rule; and the length of the turn, or the angle $\alpha/2$ subtended by it at the centre, is given by a_0 as in eq. (19). Note that the amount of rotation produced by J on the Poincaré sphere $\mathscr{P}$ is α, namely twice the length of the turn.

For the identity element, $\alpha=0$ and the turn has zero length. This is the null turn, represented by any point on S^2. For $J=-1$, we can choose $\hat{n}'=-\hat{n}$ freely on S^2, so the directed arc is a great semi circle. There is no preferred axis, and the equivalence class in this case consists of all great semi circles.

While we emphasize that a turn is an equivalence class of directed arcs, in the sequel we shall for simplicity often refer to a representative arc itself as a turn. This should cause no confusion.

The turns of an element and its inverse are simply related. Indeed, from eq. (21) it follows that

$$A(\hat{n},\hat{n}')^{-1} = A(\hat{n}',\hat{n}). \tag{22}$$

That is, the inverse of a turn is obtained by simply changing the sense of the turn, very much like the case of translation vectors. It then happens that even the composition of two turns is quite analogous (except for noncommutativity!) to the composition of translation vectors. To see this, note first that any two great circles on S^2 always intersect. Therefore given an arc on each, we can slide them on their respective great circles until the head of the first arc coincides with the tail of the second. This means that any two elements of SU(2) can be written in the form $J = A(\hat{n},\hat{n}')$, $J' = A(\hat{n}',\hat{n}'')$ with one common argument. Now it is straightforward to verify that

$$\begin{aligned} J'J &= A(\hat{n}',\hat{n}'')A(\hat{n},\hat{n}') = (\hat{n}'\cdot\hat{n}'' - i\hat{n}'_{\wedge}\hat{n}''\cdot\boldsymbol{\rho})(\hat{n}\cdot\hat{n}' - i\hat{n}_{\wedge}\hat{n}'\cdot\boldsymbol{\rho}) \\ &= \hat{n}\cdot\hat{n}'' - i\hat{n}_{\wedge}\hat{n}''\cdot\boldsymbol{\rho} = A(\hat{n},\hat{n}''). \end{aligned} \tag{23}$$

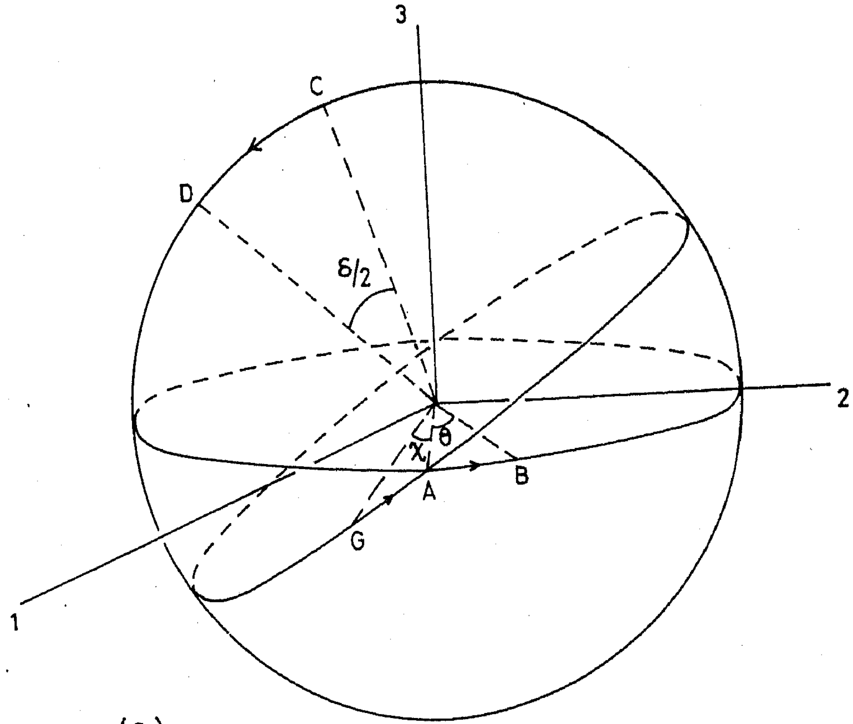
Thus the SU(2) composition law translates to turns in this way: to compose two SU(2) elements, slide the two turns on their respective great circles so that the tail of the second coincides with the head of the first; then the product of the two elements corresponds to the turn from the tail of the first turn to the head of the second. In other words, the following diagram commutes:

$$\begin{tikzcd} J, J' \in \mathrm{SU(2)} \arrow[r, "\text{SU(2) to turn rule}"'] \arrow[d, "\substack{\text{SU(2)}\\ \text{composition}\\ \text{rule}}"'] & \text{turns for } J', J \arrow[d, "\substack{\text{“addition”}\\ \text{of turns}}"] \\ J'J \in \mathrm{SU(2)} \arrow[r, "\text{SU(2) to turn rule}"'] & \text{turns for } J'J \end{tikzcd}$$

This geometric "addition" rule for turns is analogous to the parallelogram law for adding Euclidean translation vectors. The differences are that it is noncommutative; and while with vectors it suffices to translate only one of them, with turns it is in general necessary to slide both unless they commute.

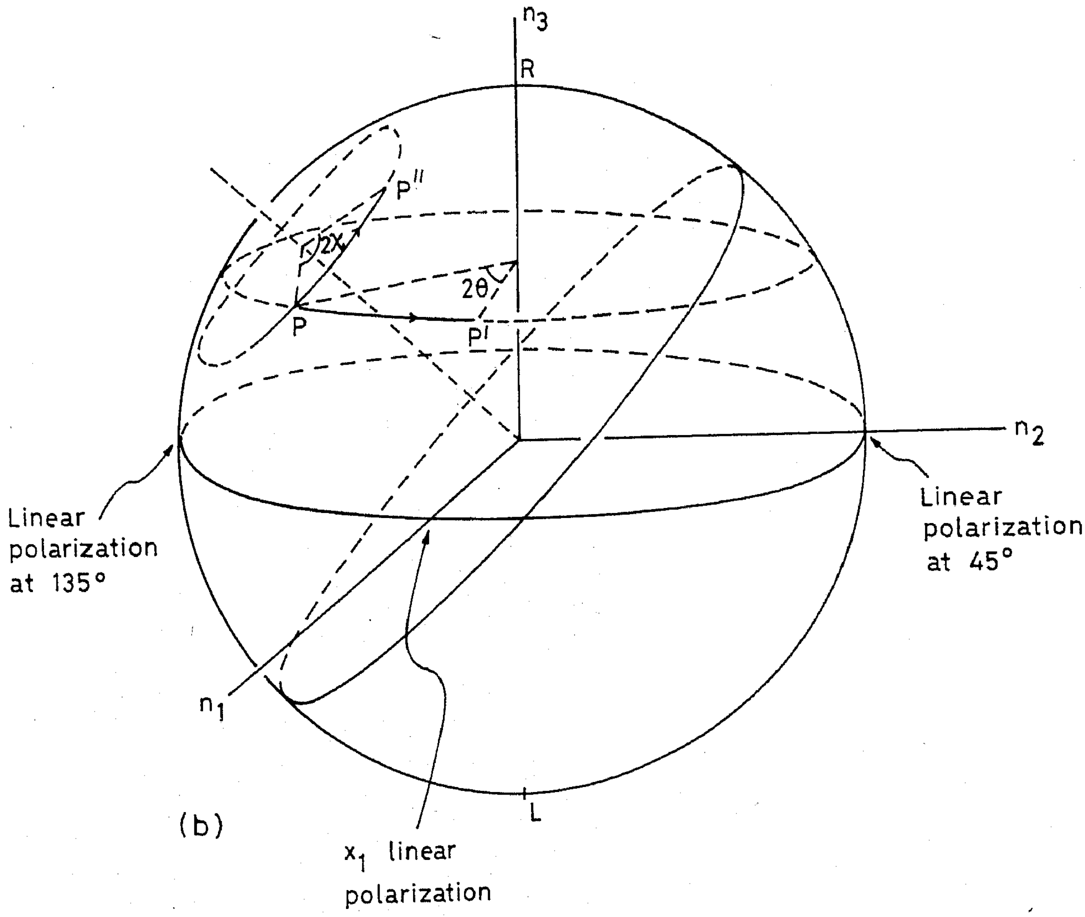
In the usual description of polarization optics the action of a system on $\mathscr{P}$ as a rotation is represented geometrically, but the resultant of a sequence of systems is computed algebraically. With the use of turns, both are describable in geometric terms.

In addition to the sphere $\mathscr{P}$ whose points represent pure polarization states, we now use another sphere $\mathscr{T}$—the "sphere of turns"—on which optical systems are represented as great circle arcs. The two spheres are shown in figure 1. The optical rotator (3), whose action on $\mathscr{P}$ has been described in §2, is represented by a great circle arc along the equator of $\mathscr{T}$, with tail at azimuth θ_0 and head at azimuth $\theta_0 + \theta$: the length of the turn is θ, and θ_0 is arbitrary. In figure 1(a), AB represents this turn.



(a)

Figure 1a. The rotator, compensator and general turns on $\mathscr{T}$.



(b)

Figure 1b. Actions of rotation and general turns on $\mathscr{P}$.

The birefringent plate $(C_\delta)_0$ of eq. (4) produces a rotation about n_1 of amount δ, on $\mathscr{P}$. So its turn is an arc on $\mathscr{T}$ in the 2–3 plane with length $\delta/2$, and sense such as to make the positive first axis normal to the plane of the turn. This is shown as CD in figure 1(a). More generally, for the system $(C_\delta)_\theta$ of eq. (5) which is a birefringent plate with fast axis at an angle θ to the x_1 axis, the arc on $\mathscr{T}$ is in the plane normal to the radial vector at azimuth 2θ in the equatorial plane, has length $\delta/2$, and sense such as to make this radial vector the positive normal. The turn GA represents a system with both birefringence and optical activity.

Given a turn on $\mathscr{T}$, the action of the corresponding optical system on a state represented by a point on $\mathscr{P}$ is straightforward. Draw the (unique) plane containing the state and parallel to the turn. This plane intersects $\mathscr{P}$ in a circle which can be visualized as the circle of constant latitude with respect to the axis of the given turn and containing the given state. Under the action of the turn, the state moves on this circle in the same sense as the turn but through twice the angle of the turn, as measured from the centre of the circle. These statements do not cover the two trivial elements $\pm\mathbb{1}\in SU(2)$ which produce the identity transformation on $\mathscr{P}$. In figure 1(b) we show how a general state $P\in\mathscr{P}$ undergoes displacements caused by various turns. The rotator turn AB carries P to P'; the general turn GA takes P to P''.

To conclude this section we note that in this new geometric representation the combined effect of a sequence of systems on the input polarization state is computed in two steps: first, the turns of the individual systems are geometrically "added" on $\mathscr{T}$ in the correct order to get the resultant turn; then the effect of this resultant on the state under consideration is again obtained geometrically by the latitude construction of the preceding paragraph.

4. Some applications of the new representation

The new geometrical representation for polarization optics, developed in the last section and based on the two spheres $\mathscr{P}$ and $\mathscr{T}$, can be used for a better understanding and sometimes faster computation of the action of a sequence of optical elements. We now present some examples to illustrate this. All the examples are chosen in the context of topological phase, which is a topic of considerable current interest. We therefore begin with some general comments on topological phases in polarization optics.

4.1 *Pancharatnam angle and Berry–Aharonov–Anandan phase*

Berry's work on geometrical phases in the adiabatic evolution of quantum mechanical states (Berry 1984, 1985, 1987a, b) has generated a great deal of interest, particularly in polarization optics. An optical realization of Berry's original ideas was reported by Chiao and Wu (1986) and Tomita and Chiao (1986). In this experiment plane polarized light is passed through an optical fibre wound in the form of an helix, and the geometrical phase manifests itself as a rotation of the plane of polarization of the output beam with respect to the input beam. Clearly, as the beam is guided along the helix the propagation direction traces a closed curve on the unit sphere S^2 of propagation directions. For a uniform helix, this curve is a circle of constant latitude with respect to the axis of the helix. It turns out that the topological phase equals the solid angle that this circuit subtends at the centre of the sphere.

Berry's topological phase was subsequently generalized by Aharonov and Anandan (1987) beyond the adiabiatic approximation. Their formulation applies to any closed circuit in the space of pure state density matrices (the ray space or the projective Hilbert space), so that all one requires for the validity of their formulation is that the physical state returns to itself after evolution over a period T(cyclic evolution). In polarization optics this corresponds to closed circuits on $\mathscr{P}$, since here E plays the role of the state vector, and Φ that of the density matrix. Given such a closed circuit in the general case, there are infinitely many different dynamical evolutions which can take the physical state over this closed circuit. The total phase picked up by the state during the cycle will be different for different evolutions, i.e., it will depend on the Hamiltonian. This is also true of the dynamical phase, defined as the time integral of the instantaneous expectation value of the Hamiltonian. But these authors showed that the difference between these two phases—total minus dynamical—is independent of the particular evolution considered and depends only on the geometry of the closed circuit. This difference is the geometrical phase of Aharonov-Anandan.

More recently Ramaseshan and Nityananda (1986) made the important observation that a phase deficiency discovered by Pancharatnam (1956) in the 50's in connection with his studies on polarized light could be considered as an early example of Berry's phase. In fact this Pancharatnam angle is more akin to the Aharonov-Anandan phase than to the Berry phase. The reason for this is easy to see. In the Pancharatnam type of situation, if x_3 is taken to be the propagation axis of the system, quantum mechanical evolution in time t is to be replaced by the unitary evolution of E in x_3, and the Hamiltonian by the generator of this evolution, as realized by the optical components. Since such polarization experiments invariably involve discrete elements, the generator of the evolution (the Hamiltonian) changes discontinuously; and one can hardly claim that the state remains an eigenstate of the x_3-dependent Hamiltonian throughout the cycle. However the Berry phase as originally formulated corresponds to situations where the evolution is adiabatic, so that the state undergoing cyclic evolution remains an eigenstate of the instantaneous Hamiltonian throughout the evolution.

To be specific, a typical polarization experiment using optical rotators and birefringent plates corresponds to unitary evolution under piecewise constant Hamiltonians. That is

$$i\frac{\partial}{\partial x_3}E(x_3)=H(x_3)E(x_3), \tag{24}$$

where the hermitian generator $H(x_3)$ is piecewise constant in x_3. Clearly, for the optical rotator (3) H is a linear combination of the identity matrix and ρ_3; for the birefringent plate (5) whose fast axis makes an angle θ with the x_1-axis, it is a linear combination of the identity matrix and $(\rho_1\cos 2\theta+\rho_2\sin 2\theta)$. For free propagation the generator is proportional to the identity matrix. In all these cases, one can appreciate that in polarization experiments the adiabaticity condition is generally not satisfied.

More recent experiments on the Pancharatnam angle, inspired by the work of Berry and of Aharonov-Anandan, have been reported (Bhandari and Samuel 1988; Chyba *et al* 1988; Simon *et al* 1988). Just as in Pancharatnam's original experiment, the set up in (Bhandari and Samuel 1988) uses a polarizer (along with two quarter-wave plates) and hence does not correspond to unitary evolution. The experiments in (Chyba *et al* 1988; Simon *et al* 1988) use, in contrast, only quarter-wave plates and hence correspond to unitary evolutions.

4.2 *The new representation and some synthesis problems*

We have already noted that a birefringent plate produces a rotation about an equatorial axis, and an optically active medium produces a rotation about the polar axis, on $\mathscr{P}$. It will be shown later, by an elementary geometrical construction on $\mathscr{T}$, that any rotation on $\mathscr{P}$ (in fact any SU(2) element) can be synthesized using an appropriate birefringent plate and an optical rotator.

In a typical experimental measurement of the Pancharatnam-Berry-Aharonov-Anandan phase one needs to change the closed circuit on $\mathscr{P}$ in such a way that this change does not affect the value of the dynamical phase, so that the consequent measured change in total phase could be identified with a change in the geometric phase alone. Now, if the polar axis rotation on $\mathscr{P}$ were realized by propagation through an optically active medium of a particular specific optical activity, an increased (decreased) rotation designed to change the circuit on $\mathscr{P}$ can only be achieved by increasing (decreasing) the length of the medium. In this process the dynamical phase is sure to change since the U(1) part of the phase is proportional to the thickness of the medium. It follows that an optically active medium is not a suitable candidate for effecting polar axis rotations in experiments designed to measure the geometric phase directly. We are faced with the problem of synthesis of a variable optical rotator with constant dynamical phase, at least for specified initial states on $\mathscr{P}$.

It turns out that two half-wave-plates can be used to produce a polar axis rotation through any angle in such a way that the dynamical phase is independent of this angle. Our new geometrical representation using $\mathscr{T}$ brings this out in a vivid pictorial manner, as shown in figure 2(a).

We wish to realize a system which has the Jones matrix of the form (3) and produces a 2θ rotation about the polar axis of $\mathscr{P}$. Clearly on $\mathscr{T}$ its turn is the equatorial arc AB in figure 2(a), with tail at azimuth θ_0 and head at azimuth $\theta_0+\theta$, for any value of θ_0. This turn has length θ. Through A and B draw the great circles passing through the poles of $\mathscr{T}$, and consider the turns AC and CB. Both are meridional arcs of length $\pi/2$, hence both correspond to half wave plates. A moment's reflection shows that the half-wave plate corresponding to AC has its fast axis at an angle $\frac{1}{2}(\theta_0+\pi/2)$, and that corresponding to CB at an angle $\frac{1}{2}(\theta_0+\theta-\pi/2)$, with respect to the x_1-axis. Now from the spherical triangle ABC it is clear that

$$(\text{turn CB})(\text{turn AC}) \equiv \text{turn AC ``+'' turn CB} = \text{turn AB}. \tag{25}$$

The noncommutative "addition" of turns has been indicated here by "+". Thus the combination of two half-wave-plates, with the fast axis of the second at an angle $\frac{1}{2}(\theta-\pi)$ with respect to that of the first, amounts identically to the optical rotator R_θ of (3). The fact that addition of turns is non-commutative manifests itself in this particular example in the following way:

$$\text{turn CB ``+'' turn AC} = -\,\text{turn AB} = (\text{turn AB})^{-1}. \tag{26}$$

One can see this result geometrically by sliding CB and AC to the northern hemisphere of $\mathscr{T}$.

It remains to be shown that if a certain state traces a closed circuit on $\mathscr{P}$ under the action of this two-half-wave-plates arrangement, the dynamical phase it picks up is independent of θ; it must of course be independent of θ_0 as well. Specifying the system components amounts to choosing the piecewise constant Hamiltonian of

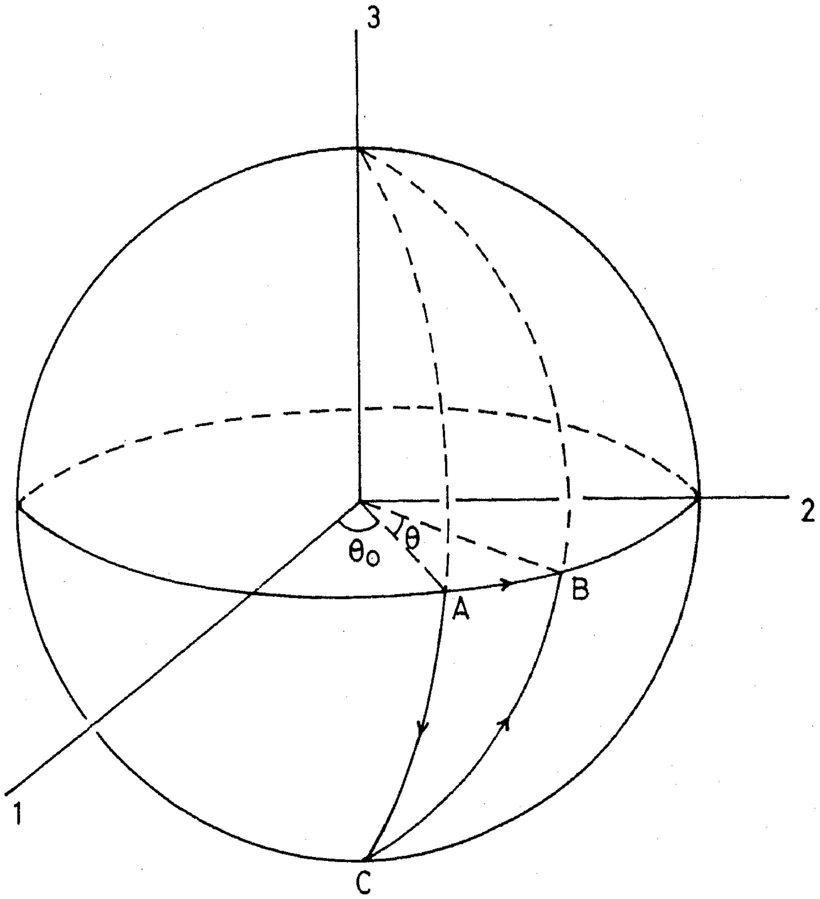


Figure 2a. Synthesis of rotator turn on $\mathscr{T}$ by half wave plates.

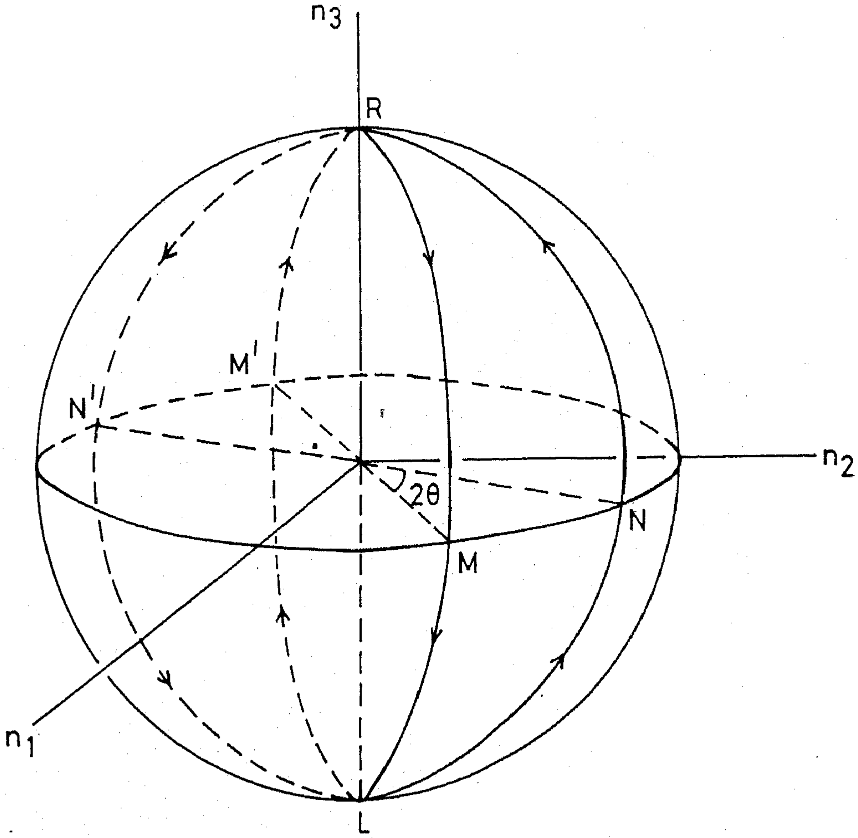


Figure 2b. Closed circuits traced on $\mathscr{P}$ under two half-wave-plate arrangement.

eq. (24) in a definite way, and therefore to a determination of the dynamical phase as well.

It is clear that the circular polarization states are the ones which trace closed circuits on $\mathscr{P}$, since they are the eigenstates of the combined transformation. As shown in figure 2(b), the state of right circular polarization traces the closed circuit RMLNR on $\mathscr{P}$. Since diametrically opposite points on $\mathscr{P}$ remain so under all rotations of $\mathscr{P}$ (i.e. orthogonal states remain orthogonal under unitary transformations), the state of left circular polarization follows the closed circuit obtained by inverting RMLNR through the origin. Thus, if M' and N' are the antipodes of M and N, the inverted circuit is LM′RN′L.

The dynamical phase is well-defined for every segment of the circuit, and the total dynamical phase is the sum of contributions from the various segments. It is easy to see that if, under the action of a *constant* Hamiltonian, a state traces a *great circle* arc on $\mathscr{P}$, then for that segment the SU(2) dynamical phase is zero. To this end recall that the dynamical phase is the x_3-integral of the expectation value of the Hamiltonian. If a constant Hamiltonian $H = \hat{a}\cdot\boldsymbol{\rho}$ has the effect of driving $\hat{n}$ in $\Phi = \frac{1}{2}(\mathbb{1} + \hat{n}\cdot\boldsymbol{\rho})$ along a great circle, then necessarily

$$\hat{a}\cdot\hat{n} = 0. \tag{27}$$

Hence the expectation value of the Hamiltonian, and the dynamical phase, are both zero:

$$E^{\dagger}HE = \operatorname{tr} H\Phi = \tfrac{1}{2}\hat{a}\cdot\hat{n} = 0. \tag{28}$$

Since the circuit in figure 2(b) consists of two great circle arcs generated by two constant Hamiltonians, it follows that the SU(2) dynamical phase is zero, and the dynamical phase is entirely due to the U(1) part of the Hamiltonian. But the U(1) phase does not depend on how the fast axes of the half-wave-plates are oriented in the $x_1 - x_2$ plane. We have thus proved the following assertion:

Two half-wave-plates, the fast axis of the second making an angle $\frac{1}{2}(\theta - \pi)$ with respect to that of the first, act as an optical rotator R_θ in such a manner that the SU(2) dynamical phase is independent of the angle θ, it being understood that the input states are of circular polarization.

Since rotating a half-wave-plate (or any birefringent plate) through π in its plane is an identity operation, we could have used $\frac{1}{2}(\theta + \pi)$ in place of $\frac{1}{2}(\theta - \pi)$ in this assertion.

It should be noted that while the turns AC, CB in figure 2(a) for the two half-wave plates do depend on θ_0, the angle θ in R_θ corresponding to the effect of the combined turn is independent of θ_0. Further, while the actual closed circuit traced on $\mathscr{P}$ under the action of the two half-wave-plates depends on θ_0, the initial state that traces this circuit, the dynamical phase and the geometric phase are all independent of θ_0. The initial state is independent of θ as well.

Generalization of these considerations to a great polygon (closed circuit formed by great circle arcs) on $\mathscr{P}$ is straightforward. First, note that evolution of a state along a plane curve on $\mathscr{P}$ can be produced by a constant Hamiltonian, and hence evolution along a great polygon by a piecewise constant Hamiltonian. The previous analysis (basically eqs (27, 28) then shows that with such a Hamiltonian the great circle arcs do not contribute to the SU(2) dynamical phase. But there remains the possibility that at one or more vertices of the polygon there act Hamiltonians for which the respective vertices form eigenstates. In such a case these pieces of evolution which do not displace the state on $\mathscr{P}$ contribute only to the dynamical phase and not to the geometric phase. If this possibility is ruled out, one can state that SU(2) evolution on a great polygon under a piecewise constant Hamiltonian contributes only to the geometric phase.

Our geometric representation also makes transparent the interesting way in which any SU(2) transformation can be synthesized using only birefringent plates. Let $J \in SU(2)$ be the Jones matrix of the system we wish to synthesize. We can slide its turn on $\mathscr{T}$ until its tail lies on the equator and its head in the northern hemisphere, as shown by AC in figure 3. Now draw the meridian through C, and let it intersect the equator at B. Both AB and BC are great circle arcs (lengths less than π), hence define turns; and the original AC is the "sum" of these:

$$\text{turn AC} = \text{turn AB "+" turn BC.} \tag{29}$$

We have already discussed the realization of the rotation turn AB using half-wave-plates, figure 2(a). The turn BC is a meridional one, representing a birefringent plate. Thus, to realize any SU(2) element one needs just a birefringent plate to follow the two half-wave-plate arrangement, with only the former depending specifically on the particular SU(2) element being synthesized.

Instead of sliding the given turn to the position AC on $\mathscr{T}$, we could equally well arrange for its tail to be in the southern hemisphere and its head on the equator, corresponding to MA in figure 3. Now let the meridian through M intersect the equator at L. Again it is clear that the turn ML followed by LA yields the given element MA. It is easily seen that ML corresponds to the same birefringent plate as

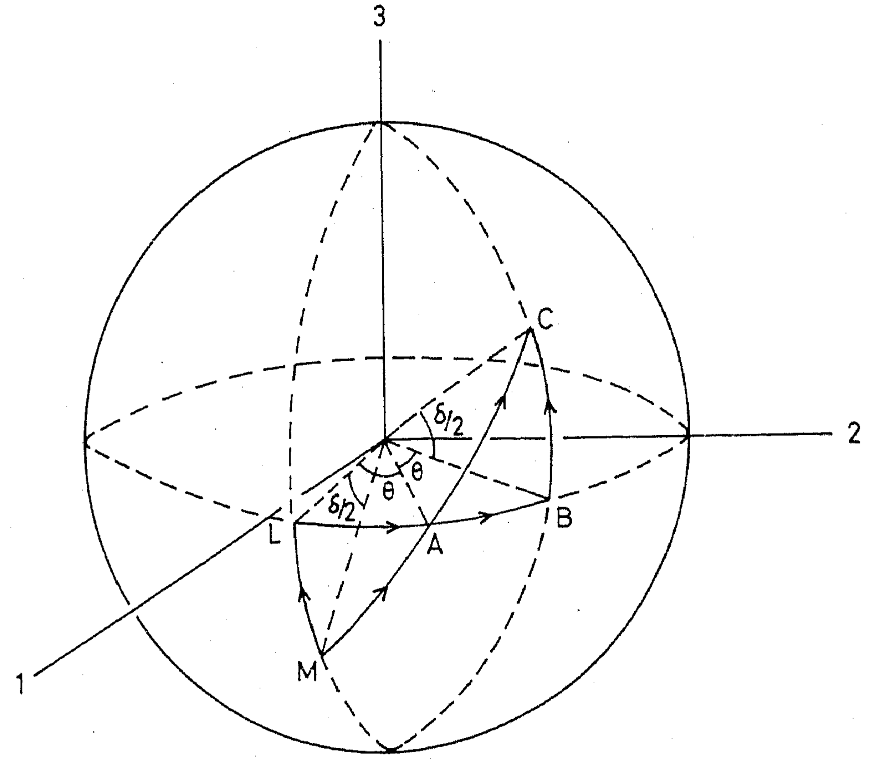


Figure 3. Synthesis of a general turn on $\mathscr{T}$.

BC, but appropriately rotated in its plane. However, LA and AB represent the same rotation turn.

This geometric description of the synthesis of any SU(2) element using only birefringent plates is a pictorial representation of the fact that any SU(2) matrix can be written as the product of an exponential of ρ_3 and an exponential of a linear combination of ρ_1 and ρ_2. (The two constructions shown in figure 3 correspond to the two possible orders of these exponentials). We maintain that while the analytic proof of this fact involves some algebra, in our geometric description it becomes obvious!

To conclude this section, we extend this all birefringent-plate synthesis of SU(2) elements one step further leading to a single gadget capable of realizing every SU(2) element. From the foregoing analysis and figure 3, it is clear that what are needed are a variable polar axis rotator and a variable birefringent plate corresponding to the turns AB and BC respectively. The variable rotator—turn AB—has been realized by the two-half-wave-plate arrangement. To realize the variable birefringent plate—turn BC—in a similar way with *fixed* optical elements, it is only necessary to see that such a turn is obtainable by conjugation within SU(2) from a rotator type turn. This is shown in figure 4. If R is a polar axis rotator (two-half-wave-plate arrangement) and Q a quarter-wave-plate with fast axis along x_1, we claim that the arrangement QRQ^{-1} realizes a variable birefringent plate. We use the method of turns to show this. In figure 4, the equatorial turn AB of length θ represents the rotator R; the point A has been chosen as $(0, -1, 0)$. Let C′N and CA, both lying in the 2-3 plane, represent

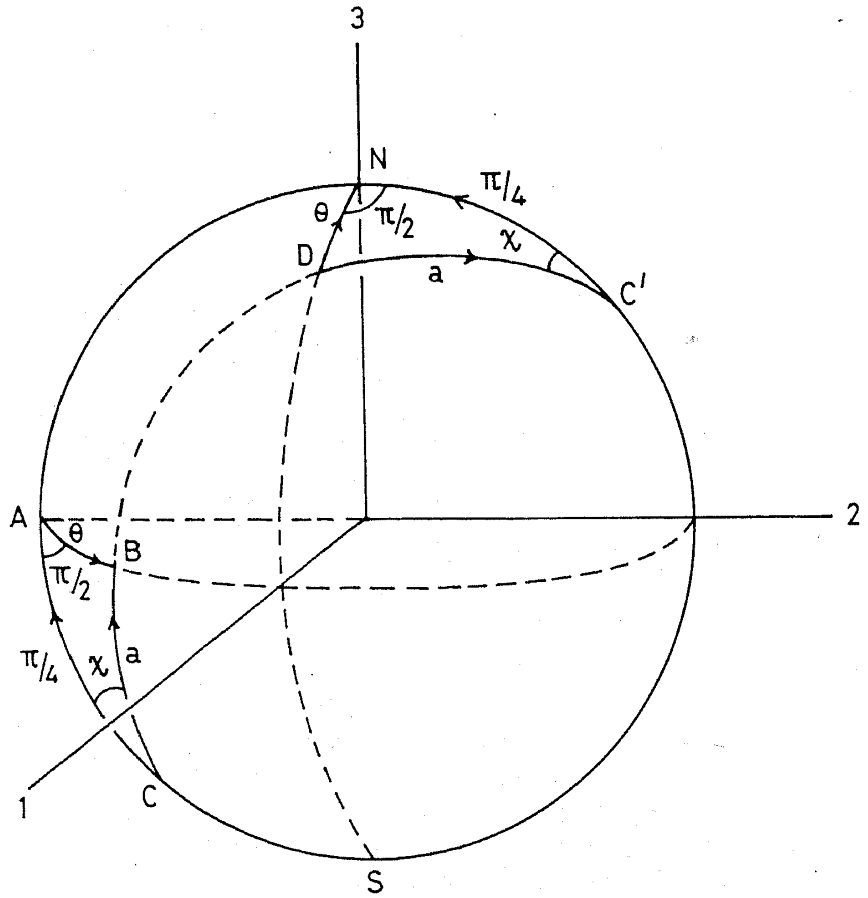


Figure 4. Relation between rotator and birefringent plate.

Q and Q^{-1} respectively; each is of length $\pi/4$. From the spherical triangle ABC it is clear that CB is the turn representing RQ^{-1}. Let its length be a, and let χ be the angle between CA and CB. Slide CB to the equivalent position DC′. We will see in a moment that D lies in the 1–3 plane. The great circle NACSC′N (in the 2–3 plane) and the one containing C′DBC, making an angle χ between them, intersect at C and C′. Hence the angle between DC′ and C′N is χ. Since DC′ has length a while C′N has length $\pi/4$, we see that apart from orientation the spherical triangles DC′N and BCA are similar. Hence the angle DNC′ is $\pi/2$, and DN has the same length θ as AB. But as DC′ is the turn for RQ^{-1} and C′N is that for Q, we find that DN represents QRQ^{-1}. It is a birefringent plate with fast axis at $\pi/4$ to the x_1 axis, and with $\delta = 2\theta$. Since θ which gives the relative orientation of the two half-wave-plates in R can be varied, we see that QRQ^{-1} synthesizes a variable birefringent plate with fast axis at $\pi/4$ to the x_1 axis. To change the fast axis, one has to simply rotate QRQ^{-1} about the beam axis (x_3-axis).

To summarize, the combination of a two-half-wave-plate arrangement R′ followed by QRQ^{-1}, i.e. $QRQ^{-1}R'$, realizes every SU(2) element. The part R′ with variable θ' (the angle between the fast axes of the two-half-plates in R′ being $\frac{1}{2}(\theta' + \pi)$) realizes the turn AB in figure 3, and has length θ'. The part QRQ^{-1} realizes BC: the angle θ in R is the length of BC, and rotation of QRQ^{-1} (or equivalently of $QRQ^{-1}R'$ as a whole) chooses the meridian on which BC falls. The gadget consists on the whole of six birefringent plates, four being half-wave-plates and the other two being quarter-wave-plates. The variable parameters are the relevant angle between the fast axes in each half-wave-plate pair, and the overall orientation of the system (transverse to the beam axis), this accounting for the three parameters of SU(2).

5. The geometric phase as a solid angle

We have earlier recalled the important result of Aharonov-Anandan (1987): when a state evolves over a closed circuit in the projective Hilbert space, the difference between the total phase picked up by the state and the dynamical phase (the time integral of the instantaneous expectation value of the Hamiltonian) is independent of the particular Hamiltonian and depends on the circuit alone. We shall show in this section that for SU(2) systems for which the projective state space is the Poincaré sphere $\mathscr{P}$, given any closed circuit on $\mathscr{P}$ this phase difference is half the solid angle subtended by the circuit at the centre of $\mathscr{P}$. The proof is structured in such a way that it easily generalizes to the noncompact group $SU(1,1) \simeq SP(2,R) \simeq SL(2,R)$ which plays an important role in first order Fourier optics and the problem of squeezed states.

Given $\hat{n}$, $\hat{n}' \in S^2$, we need to construct a distinguished SU(2) element, $B(\hat{n}, \hat{n}')$ say, which rotates $\hat{n}$ to $\hat{n}'$ along the (shorter) great circle arc. The element $A(\hat{n}, \hat{n}') \in SU(2)$ in eq. (21) will not do this, since it was tailored to obtain the composition law (23). In fact, $A(\hat{n}, \hat{n}')$ rotates $\hat{n}$ beyond $\hat{n}'$ to $\hat{n}''$ in the plane of $\hat{n}$ and $\hat{n}'$, such that the angle between $\hat{n}$ and $\hat{n}'$ equals that between $\hat{n}'$ and $\hat{n}''$:

$$
\begin{aligned}
&A(\hat{n}, \hat{n}')\hat{n}\cdot\boldsymbol{\rho} A(\hat{n}, \hat{n}')^{-1} = \hat{n}''\cdot\boldsymbol{\rho}, \\
&\hat{n}'' = 2\hat{n}\cdot\hat{n}'\hat{n}' - \hat{n}, \\
&\hat{n}\cdot\hat{n}' = \hat{n}'\cdot\hat{n}'', \quad \hat{n}\wedge\hat{n}' = \hat{n}'\wedge\hat{n}''.
\end{aligned} \tag{30}
$$

On $\hat{n}'$ the effect is similar:

$$A(\hat{n},\hat{n}')\hat{n}'\cdot\boldsymbol{\rho}A(\hat{n},\hat{n}')^{-1} = (2\hat{n}'\cdot\hat{n}''\hat{n}'' - \hat{n}')\cdot\boldsymbol{\rho}. \tag{31}$$

But these properties of $A(\hat{n},\hat{n}')$ show that the element $B(\hat{n},\hat{n}')$ we are in need of, and which will take $\hat{n}'$ while leaving $\hat{n}_\wedge\hat{n}'$ invariant, can be obtained by replacing $\hat{n}'$ in $A(\hat{n},\hat{n}')$ by the unit bisector of $\hat{n}$ and $\hat{n}'$:

$$\begin{aligned} B(\hat{n},\hat{n}') &= A(\hat{n},(\hat{n}+\hat{n}')/|\hat{n}+\hat{n}'|) \\ &= [2(1+\hat{n}\cdot\hat{n}')]^{-1/2}(\mathbb{1} + A(\hat{n},\hat{n}')), \\ B(\hat{n},\hat{n}')\hat{n}\cdot\boldsymbol{\rho}\, B(\hat{n},\hat{n}')^{-1} &= \hat{n}'\cdot\boldsymbol{\rho}, \\ A(\hat{n},\hat{n}') &= (\mathrm{sign}\,(\hat{n},\hat{n}'))B(\hat{n},\hat{n}''). \end{aligned} \tag{32}$$

(Here in the last line $\hat{n}''$ is as in eq. (30)). As one should expect, this expression for $B(\hat{n},\hat{n}')$ becomes indeterminate when $\hat{n}\cdot\hat{n}' = -1$; the two unit vectors then have a one-parameter family of bisectors rather than a unique one, or equivalently a one-parameter family of great circle arcs connecting them.

For $\hat{n}$, $\hat{n}'\in S^2$, the most general constant Hamiltonian which generates a unitary transformation taking $\hat{n}$ to $\hat{n}'$ is not unique, but is an arbitrary linear combination of $\hat{n}_\wedge\hat{n}'\cdot\boldsymbol{\rho}$ and $(\hat{n}+\hat{n}')\cdot\boldsymbol{\rho}$. (This nonuniqueness is over and above the reparametrization of the evolution parameter resulting in a scale change of the Hamiltonian, which freedom is always present and is ignored). A Hamiltonian proportional to $\hat{n}_\wedge\hat{n}'\cdot\boldsymbol{\rho}$ alone takes $\hat{n}$ to $\hat{n}'$ along a great circle arc; one proportional to $(\hat{n}+\hat{n}')\cdot\boldsymbol{\rho}$ does so along a circle of constant latitude with respect to the axis $\hat{n}+\hat{n}'$, which incidentally is the smallest circle containing $\hat{n}$ and $\hat{n}'$. Under each constant Hamiltonian $\hat{n}$ traces a plane curve on $\mathscr{P}$; conversely for each plane curve connecting $\hat{n}$ to $\hat{n}'$, we have a unique constant Hamiltonian that drives $\hat{n}$ along it. Both $A(\hat{n},\hat{n}')$ and $B(\hat{n},\hat{n}')$ are generated by $\hat{n}_\wedge\hat{n}'\cdot\boldsymbol{\rho}$.

We are now equipped to carry out the calculation of the geometric phase. For $0\leqslant s\leqslant 1$, let $\hat{n}(s)$ be an arbitrary smooth closed circuit Γ on $\mathscr{P}$, as shown in figure 5. Let $\hat{n}(0) = \hat{n}(1) = \hat{n}_0$. As s varies, the state

$$\Phi(s) = \tfrac{1}{2}(\mathbb{1} + \hat{n}(s)\cdot\boldsymbol{\rho}) \tag{33}$$

evolves along Γ, starting and ending at $\Phi(0) = \Phi(1)$. Let $J(s)\in SU(2)$ for $0\leqslant s\leqslant 1$ be the most general choice of unitary operators which produce this evolution:

$$\Phi(s) = J(s)\Phi(0)J(s)^{-1}. \tag{34}$$

We recover the s-dependent Hamiltonian $H(s)$ generating this evolution through

$$\begin{aligned} &\frac{\mathrm{d}J(s)}{\mathrm{d}s} \equiv \dot{J}(s) = -iH(s)J(s), \\ &J(0) = \mathbb{1}, \\ &J(s) = \mathscr{P}\exp\left(-i\int_0^s H(s')\,\mathrm{d}s'\right). \end{aligned} \tag{35}$$

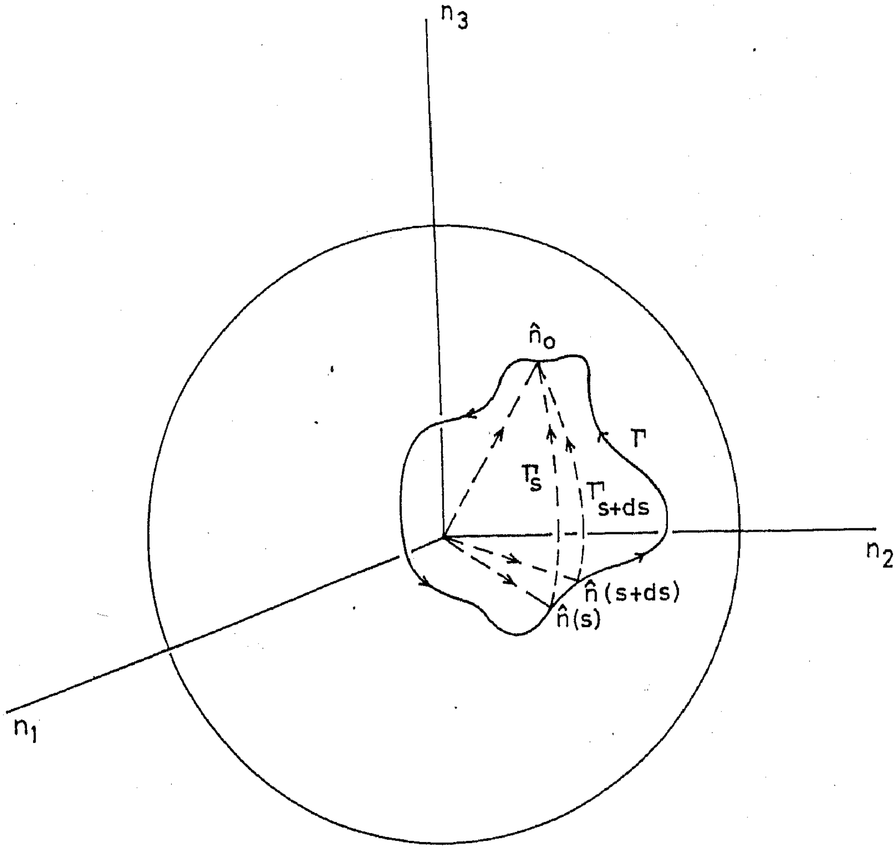


Figure 5. Calculating the geometric phase on $\mathscr{P}$.

From our considerations in the previous paragraph it follows that

$$H(s) = \tfrac{1}{2}\hat{n}(s)_{\wedge}\dot{\hat{n}}(s)\cdot\boldsymbol{\rho} + \alpha(s)\hat{n}(s)\cdot\boldsymbol{\rho}, \tag{36}$$

where $\alpha(s)$ is an arbitrary real c-number function of s. Since $\hat{n}(s)\cdot\dot{\hat{n}}(s) = 0$, we can readily check that

$$\frac{\mathrm{d}}{\mathrm{d}s}(J(s)^{-1}\hat{n}(s)\cdot\boldsymbol{\rho}J(s)) = J(s)^{-1}(i[H(s), \hat{n}(s)\cdot\boldsymbol{\rho}] + \dot{\hat{n}}(s)\cdot\boldsymbol{\rho})J(s) = 0, \tag{37}$$

showing that $J(s)$ indeed produces the required evolution along Γ:

$$J(s)\hat{n}_0\cdot\boldsymbol{\rho}J(s)^{-1} = \hat{n}(s)\cdot\boldsymbol{\rho}. \tag{38}$$

Now for each s in the range $0 \leqslant s \leqslant 1$ consider the s-dependent closed circuit Γ_s consisting of the portion of Γ from $\hat{n}_0$ to $\hat{n}(s)$ followed by the great circle arc from $\hat{n}(s)$ back to $\hat{n}_0$. For $s = 1$, Γ_1 coincides with the original Γ. We may choose the closed circuit evolution over Γ_s to be produced by

$$Q(s) = B(\hat{n}(s), \hat{n}_0)J(s),$$
$$Q(0) = 1, \quad Q(1) = J(1). \tag{39}$$

By construction $Q(s)$ for each s returns $\hat{n}_0$ to itself, hence $\Phi(0)$ is an eigenstate of $Q(s)$. Thus $Q(s)$ must necessarily be of the form

$$Q(s) = \exp(-i\varphi(s)\hat{n}_0\cdot\boldsymbol{\rho}),$$
$$\varphi(0) = 0, \tag{40}$$

where the real c-number function $\varphi(s)$ is the total phase picked up by the state $\hat{n}_0$ on evolution over the closed circuit Γ_s. We are interested in $\varphi(1)$, the total phase for the circuit Γ and the Hamiltonian (36).

The problem has been reduced to evaluation of $\varphi(s)$. We develop a first order differential equation for it. From eqs (32, 39), we have:

$$\dot{Q}(s) = \dot{B}(\hat{n}(s),\hat{n}_0)J(s) + B(\hat{n}(s),\hat{n}_0)\dot{J}(s),$$
$$\dot{B}(\hat{n}(s),\hat{n}_0) = \frac{1}{\sqrt{2}}\left[\frac{-1}{2}(1+\hat{n}(s)\cdot\hat{n}_0)^{-3/2}\hat{n}_0\cdot\dot{\hat{n}}(s)(\mathbb{1} + A(\hat{n}(s),\hat{n}_0)) + (1+\hat{n}(s)\cdot\hat{n}_0)^{-1/2}(\dot{\hat{n}}(s)\cdot\hat{n}_0 - i\dot{\hat{n}}(s)_\wedge\hat{n}_0\cdot\boldsymbol{\rho})\right]. \tag{41}$$

After straightforward algebra, this simplifies to

$$\dot{Q}(s)J(s)^{-1} = \dot{B}(\hat{n}(s),\hat{n}_0) - iB(\hat{n}(s),\hat{n}_0)H(s)$$
$$= \frac{-i}{2\sqrt{2}}(1+\hat{n}(s)\cdot\hat{n}_0)^{-3/2}\hat{n}_0\cdot\hat{n}(s)_\wedge\dot{\hat{n}}(s)(\hat{n}_0+\hat{n}(s))\cdot\boldsymbol{\rho}$$
$$- i\alpha(s)B(\hat{n}(s),\hat{n}_0)\hat{n}(s)\cdot\boldsymbol{\rho}. \tag{42}$$

Hence

$$\dot{Q}(s)Q(s)^{-1} = \dot{Q}(s)J(s)^{-1}B(\hat{n}(s),\hat{n}_0)^{-1}$$
$$= \frac{-i\,\hat{n}_0\cdot\hat{n}(s)_\wedge\dot{\hat{n}}(s)}{2\,(1+\hat{n}_0\cdot\hat{n}(s))}\hat{n}_0\cdot\boldsymbol{\rho} - i\alpha(s)\hat{n}_0\cdot\boldsymbol{\rho}, \tag{43}$$

which in view of eq. (40) means

$$\dot{\varphi}(s) = \frac{1}{2}\frac{\hat{n}_0\cdot\hat{n}(s)_\wedge\dot{\hat{n}}(s)}{(1+\hat{n}_0\cdot\hat{n}(s))} + \alpha(s). \tag{44}$$

Now, if $\Omega(s)$ is the solid angle subtended by the closed circuit Γ_s at the origin, then $(1+\hat{n}_0\cdot\hat{n}(s))^{-1}\hat{n}_0\cdot\hat{n}(s)_\wedge\dot{\hat{n}}(s)$ is the rate $\mathrm{d}\Omega(s)/\mathrm{d}s$ at which the great circle arc from $\hat{n}_0$ to $\hat{n}(s)$ sweeps out solid angle as $\hat{n}(s)$ moves along Γ. Further, $\alpha(s)$ is the expectation value of the Hamiltonian $H(s)$ in the instantaneous state $\Phi(s)$:

$$\langle H(s)\rangle \equiv \mathrm{tr}\, H(s)\Phi(s) = \alpha(s). \tag{45}$$

Thus we may rewrite eq. (44) as

$$\frac{\mathrm{d}\varphi(s)}{\mathrm{d}s} = \frac{1}{2}\frac{\mathrm{d}\Omega(s)}{\mathrm{d}s} + \langle H(s)\rangle, \tag{46}$$

and on integration from $s=0$ to $s=1$ we get an expression for the geometric phase:

$$\varphi(1)-\int_0^1 \mathrm{d}s\langle H(s)\rangle=\tfrac{1}{2}\Omega. \tag{47}$$

Here Ω is the solid angle subtended by the closed circuit Γ at the origin. This completes the proof of the assertion made at the start of this Section.

A difference in signature between our definition of the geometric phase and that of Aharonov-Anandan should be noted. Assuming that in both conventions the geometric phase is half the solid angle, our definition implies that the solid angle is positive (negative) for a positively (negatively) traversed circuit, while that of Aharonov-Anandan amounts to taking the solid angle positive for a negatively (clockwise) traversed circuit.

There is an interesting way in which the relation (47) between the geometric phase and the solid angle on $\mathscr{P}$ can be dynamically exhibited in a polarization experiment. Consider right circularly polarized light passing through the two-half-wave-plate arrangement discussed earlier in connection with figure 2(a, b). Upon passage through the first half-wave-plate, the input state R is transformed into the left circularly polarized state L along the meridian RML. Upon passage through the second plate L is transformed back to R through the meridian LNR. Thus the initial state has traversed the closed circuit RMLNR, which subtends the solid angle 4θ (assuming that the fast axis of the second plate is at $\theta \pm \pi/2$ with respect to that of the first plate). Hence from eq. (47) the geometric phase change suffered by the state is 2θ. If the dynamical phase is φ_0, we have shown in §4 that for this arrangement φ_0 is independent of θ.

Now suppose the second half wave plate is rotated in its plane with angular frequency ω, so that θ varies as

$$\theta(t)=\theta_0+\omega t, \tag{48}$$

where the initial value θ_0 will turn out to be irrelevant in the present situation. With this rotation, the meridian LNR in figure 2(b) rotates about the n_3 axis with angular frequency 2ω, while the meridian RML stays constant. Thus we have a closed circuit at each time, with a time-dependent solid angle

$$\Omega(t)=4\theta_0+4\omega t, \tag{49}$$

a time-dependent geometric phase

$$\beta(t)=\tfrac{1}{2}\Omega(t)=2\theta_0+2\omega t, \tag{50}$$

and a total phase $\varphi_0+\beta(t)$. Since this phase change is linear in time, it results in a shift in the frequency of the light. If the input frequency is ω_0, the output frequency will be

$$\omega'=\omega_0+2\omega. \tag{51}$$

If this output light is superposed on a reference beam extracted from the input which has frequency ω_0, then beats will be seen at the difference frequency 2ω. Such an experiment where this effect can be used to fine tune the frequency of a laser beam has been reported recently (Simon *et al* 1988). Instead of using two half-wave-plates under a single pass configuration we could equally well realize the closed circuit in figure 2(b) by using two quarter-wave-plates in a double pass configuration as in a Michelson interferometer, with the input state being a linearly polarized one corresponding to the point M in figure 2(b). This is the option actually adopted in (Simon *et al* 1988).

The above example involves a geometric phase which grows linearly in time, leading to a simple frequency shift. There are physical situations where the geometric phase could have a more complicated time dependence, leading to a general phase modulation and frequency spread. A specific situation where this could take place is explored in (Simon and Kumar 1988)

Returning to the above experiment, if we had used left circularly polarized light as input, then the closed circuit would have been LM'RN'L. While the solid angle has the same magnitude as for RMLNR, the circuits are traversed in opposite senses and hence give equal and opposite geometric phases. It follows that for left circularly polarized input light and the second half-wave-plate rotating at an angular velocity ω, the frequency shift will be in the opposite direction so that the output frequency is $\omega_0 - 2\omega$.

6. Concluding remarks

We have presented in this paper a new geometrical representation for polarization optics based on Hamilton's turns. In this representation not only the action of individual optical systems on polarization states but also the sequential composition of several systems is handled geometrically rather than algebraically. We have also presented several applications of the new representation in the context of the geometric phase, and the suggestiveness of the formalism in synthesis problems must be evident.

For simplicity we limited our attention to pure polarization states. The generalization to transformation of mixed states by unitary systems is straightforward. A mixed state can be written as an incoherent superposition of characteristic orthogonal pure states:

$$\Phi = \alpha\Phi_1 + (1-\alpha)\Phi_2. \tag{52}$$

Here Φ_1 and Φ_2 are the normalized orthogonal eigenstates of the given mixed state Φ, and without loss of generality we can assume $\frac{1}{2} \leqslant \alpha \leqslant 1$. The representation (52) is unique except in the degenerate case $\alpha = \frac{1}{2}$ when Φ becomes a multiple of the identity matrix. Since Φ_1 fixes Φ_2 uniquely, a mixed state can be represented by the pair (α, Φ_1), where Φ_1 is a point on $\mathscr{P}$ and $\frac{1}{2} \leqslant \alpha \leqslant 1$. Under the action of a unitary system Φ transforms as in eq. (15). It follows that α remains invariant, while the transformation of Φ_1 on $\mathscr{P}$ follows the analysis in §§2, 3. However, the generalization of the notion of geometric phase to mixed states has never been clear.

While our analysis has been in the specific context of polarization optics, it should be clear that all our considerations apply faithfully to quantum mechanical spin half and pseudo spin half systems. The role of the polarization matrix Φ is then played by the density matrix and that of the unit vector $\hat{n}$ on $\mathscr{P}$ by the (pseudo) spin vector. Again the geometric phase is half the solid angle subtended by the closed circuit traced by the tip of the (pseudo) spin vector at the centre of the sphere of (pseudo) spin vectors.

While the group SU(2) is compact, the closely related noncompact group SU(1, 1) plays an important role in several physical problems. We note that SU(1, 1) is a two-fold covering of the three-dimensional Lorentz group SO(2, 1), and is also the group of linear canonical transformations in one canonical pair of variables; equally well, of Bogoliubov transformations (squeezing being a subset of such transformations) for a creation-annihilation operator pair in view of the isomorphisms SU(1, 1) = SL(2, R) = SP (2, R). This group plays a fundamental role in first order Fourier optics involving axially symmetric systems and also in problems of squeezed states. Our presentation of Hamilton's turns for SU(2) was designed to make the generalization to SU(1, 1) possible. Indeed such a generalization has been recently accomplished in (Simon *et al* 1988), where the term "screws" has been used in place of turns for SU(1, 1). The role of the sphere of turns $\mathscr{T}$ is now played by the single sheeted unit hyperboloid $x_1^2 + x_2^2 - x_0^2 = 1$, and that of great circles by the intersection of planes passing through the origin with this hyperboloid. We can then employ screws to obtain a new geometric representation for first order optics and the squeezed state problem. A generalization of the geometric phase-solid angle connection to the noncompact SU(1, 1) case is also suggested.

We note that the present analysis was restricted to polarization states of plane waves and their transformation by non-image forming unitary systems. Recently we have developed a systematic method, based on the Poincaré invariance of Maxwell's equations, for handling polarization of profiled beams and their transformation by first order image forming systems (Mukunda *et al* 1983; 1985). In situations when the lenses are made of nonbirefringent non-optically active material, it has been shown that the transformations of the transverse position dependent polarization is described by the SL(2, R) group of first order optics. If the lenses are made of birefringent and/or optically active material, then the transformation of the polarization states and also the geometric phase will be expected to be governed by the product group $SL(2, R) \otimes SU(2)$.

Finally we note that, as has been shown elsewhere, every SL(2, R) first order system can be synthesized with at most three lenses (Sudarshan *et al* 1985). The powers of these lenses, as also their separations, do of course depend on the particular SL(2, R) system being synthesized. Our design in section 4 of a single gadget capable of realising all SU(2) transformations through simple rotations of the components without having to change the components themselves raises the following question: Is it possible to design a lens system which can realize every SL(2, R) element by merely adjusting the locations of the lenses in the gadget without changing their powers?

We plan to return to these various questions elsewhere.

Acknowledgement

This work was supported in part by the Department of Energy under Grant No. DE-FG05-85 ER 40200.

References

Aharonov Y and Anandan J 1987 *Phys. Rev. Lett.* **58** 1593
Berry M V 1984 *Proc. R. Soc.* (*London*) **A392** 45

Berry M V 1985 *J. Phys.* **A18** 15
Berry M V 1987 *Nature* (*London*) **326** 277
Berry M V 1987 *Proc. R. Soc.* (*London*) **A414** 31
Bhandari R and Samuel J 1988 *Phys. Rev. Lett.* **60** 1211
Biedenharn L C and Louck J D 1981 *Angular momentum in quantum physics*, Encyclopedia of mathematics and its applications (Reading, Massachusetts: Addision-Wesley Publishing Co.) Vol. 8
Born M and Wolf E 1965 *Principles of optics* (Oxford: Pergamon Press) p. 31
Chiao R Y and Wu Y S 1986 *Phys. Rev. Lett.* **57** 933
Chyba T H, Wang L J, Mandel L and Simon R 1988 Measurement of the Pancharatnam phase for a light beam, *Opt. Lett.* (in press)
Hamilton W 1853 Lectures on quaternions, Dublin
Mukunda N, Simon R and Sundarshan E C G 1983 *Phys. Rev.* **A28** 2933
Mukunda N, Simon R and Sudarshan E C G 1985 *J. Opt. Soc. Am.* **A2** 416
Poincaré H 1892 *Theorie Mathematique de la Lumiere*, (ed.) George Carré, Paris, p. 275
Pancharatnam S 1956 *Proc. Indian Acad. Soc.* **A44** 247 reprinted in Collected works of S Pancharatnam, Oxford University Press, 1975
Ramachandran G N and Ramaseshan S 1961 *Crystal optics, Handbuch der Physik*, Vol. XXV/1 (Berlin: Springer-Verlag)
Ramaseshan S and Nityananda R 1986 *Curr. Sci.* **55** 1225
Simon R, Kimble H J and Sudarshan E C G 1988 *Phys. Rev. Lett.* **61** 19
Simon R and Kumar N 1988 *J. Phys.* A (in press)
Simon R, Sudarshan E C G and Mukunda N 1989 *Phys. Rev. Lett.* **62** 1331
Sudarshan E C G, Mukunda N and Simon R 1985 *Optica Acta* **32** 855
Stokes G G 1852 *Trans. Camb. Phil. Soc.* **9** 399