# 2

# The Spin-Vector Calculus of Polarization

Spin-vector calculus is a powerful tool for representing linear, unitary transformations in Stokes space. Spin-vector calculus attains a high degree of abstraction because rules for vector operations in Stokes space are expressed in vector form; there is no *a priori* reference to an underlying coordinate system. Absence of the underlying coordinate system allows for an elegant, compact calculus well suited for polarization studies.

Spin-vector calculus is well known in quantum mechanics, especially relating to quantized angular momentum. Aso, Frigo, Gisin, and Gordon and Kogelnik have greatly assisted the optical engineering community by adopting this calculus to telecommunications applications [1, 3–5]. The purpose of this chapter is to bring together a complete description of the calculus as found in a variety of disparate sources [2, 3, 5–8], and to tailor the presentation with a vocabulary familiar to the electrical engineer. Tables 2.2 and 2.3 located at the end of the chapter offer a summary of the principal relations.

## 2.1 Motivation

The purpose of this calculus is to build a geometric interpretation of polarization transformations. The geometric interpretation of polarization states was already developed in §1.4. The Jones matrix, while a direct consequence of Maxwell's equations when light travels through a medium, is a complex-valued $2 \times 2$ matrix. This is hard to visualize. The Mueller matrix, however, can be visualized as rotations and length-changes in Stokes space. The spin-vector formalism makes a bilateral connection between the Jones and Mueller matrices.

Of all the possible Jones matrices, two classes predominate in polarization optics: the unitary matrix and the Hermitian matrix. The unitary matrix preserves lengths and imparts a rotation in Stokes space. A retardation plate is described as a unitary matrix. The Hermitian matrix comes from a measurement, such as that of a polarization state. Since all measured values must be

real quantities, the eigenvalues of a Hermitian matrix are real. The projection induced by a polarizer is described as a Hermitian matrix.

Based on the characteristic form (1.4.27) on page 19 of the Mueller matrix for a unitary matrix, defined by $UU^\dagger = I$, one can write

$$\mathbf{J}_U \longrightarrow \mathbf{M}_U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbf{R} & \\ 0 & & & \end{pmatrix} \tag{2.1.1}$$

where $\mathbf{R}$ is a $3 \times 3$ rotation matrix having real-valued entries. Since the polarization transformation through multiple media is described as the product of Jones matrices, one would expect a one-to-one correspondence between multiple unitary matrices and multiple rotation matrices. This would lead to

$$\mathbf{J}_{U_2U_1} \longrightarrow \mathbf{M}_{U_2U_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbf{R}_2\mathbf{R}_1 & \\ 0 & & & \end{pmatrix} \tag{2.1.2}$$

This is indeed the case. Moreover, the Mueller matrix representing passage of light through any number of retardation plates always keeps the form of (2.1.1). Rotation matrix $\mathbf{R}$ is therefore a group closed under rotation.

Taking the abstraction one step further, any rotation has an axis of rotation and an angle through which the system rotates. Instead of describing the rotation $\mathbf{R}$ as a $3 \times 3$ matrix, it is more general to describe the rotation as a vector quantity: $\mathbf{R} = f(\hat{r}, \varphi)$, where $\hat{r}$ is the rotation axis in Stokes space and $\varphi$ is the angle of rotation. The vector $\hat{r}$ need not be resolved onto an orthonormal basis to give $\hat{r} = \hat{x}\, r_x + \hat{y}\, r_y + \hat{z}\, r_z$; this operation may be postponed indefinitely. This is in contrast to writing $\mathbf{R}$ as a $3 \times 3$ matrix where the underlying orthonormal basis is explicit. Accordingly, $\hat{r}$ exists as a vector in vector space and can undergo operations such as rotation, inner product, and cross product with respect to other vectors.

In parallel to the unitary-matrix case, the Mueller matrix that corresponds to a Hermitian matrix, defined by $H = H^\dagger$, one can write

$$\mathbf{J}_H \longrightarrow \mathbf{M}_H = \begin{pmatrix} & & \\ & \tilde{\mathbf{H}} & \\ & & \end{pmatrix} \tag{2.1.3}$$

This indeed is a tautology. As with the unitary matrices, products of Hermitian matrices in Jones space result in products of Mueller matrices in Stokes space. That is,

$$\mathbf{J}_{H_2H_1} \longrightarrow \mathbf{M}_{H_2H_1} = \mathbf{M}_{H_2}\mathbf{M}_{H_1} \tag{2.1.4}$$

All Hermitian operations are closed within the $4 \times 4$ Mueller matrix.

As it appears, products of unitary-corresponding Mueller matrices change only entries in the lower-right-hand $3 \times 3$ sub-matrix. Inclusion of even a single Hermitian-corresponding Mueller matrix scatters those nine elements into all sixteen matrix positions. This is a non-reversible process. There is, however, a remarkable exception. A traceless Hermitian matrix $\underline{H}$, defined by $\mathrm{Tr}\underline{H} = 0$, has a corresponding Mueller matrix of the form

$$\mathbf{J}_{\underline{H}} \longrightarrow \mathbf{M}_{\underline{H}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbf{V} & \\ 0 & & & \end{pmatrix} \tag{2.1.5}$$

where $\mathbf{V}$ is a Stokes-space vector having a length and pointing direction. (Note that a rotation operator has unit length, two angles that determine the vector direction, and one angle of rotation. A Stokes vector has a length and two angles that determine the vector direction. Both rotation operator and Stokes vector have three parameters). Arbitrary products of unitary matrix $U$ and traceless Hermitian matrix $\underline{H}$ form an extended closed group in which entries change only in the lower right-hand $3 \times 3$ sub-matrix of $\mathbf{M}$.

Throughout this chapter and the chapters on polarization-mode dispersion, one looks for zero trace of Hermitian matrices. If this property is established, then a calculus that includes lossless rotations of vectors can be applied to the system. This calculus is called spin-vector calculus, and is the topic of the present chapter.

## 2.2 Vectors, Length, and Direction

Physical systems can often be described by the state the system is in at a particular time and position. The span of all possible states for a given system is called a state space. Any particular state represents all the information that one can know about the system at that time and position. Interaction between a physical system and external influences, such as transmission through media or applied force, can change the state. So, there are two categories of study: the description of state, and the transformation of state.

A state that describes wave motion can be represented by a vector with complex scalar entries. The dimensionality of the state vector is determined by the number of states that are invariant to an external influence. That is, the dimension of a state vector equals the number of eigenstates of the system. For polarization, the dimensionality is two. The important properties of a vector space are direction, length, and relative angles. These metrics will form a common theme throughout the following development.

### 2.2.1 Bra and Ket Vectors

Bra and ket spaces are two equivalent vector spaces that describe the same state space. Bra and ket spaces, or "bracket" space, is a formulation developed

by P. A. M. Dirac and used extensively in quantum mechanics. Bras and kets are vectors with dimension equal to the state dimension. When a bra space and ket space describe the same state vector, the bra and ket are duals of one another. For a state vector $\vec{a}$, the ket is written $|a\rangle$ and the bra is written $\langle a\,|$. The entries in bra and ket vectors are complex scalar numbers. A ket vector suitable for polarization studies is

$$|a\rangle = \begin{pmatrix} a_x \\ a_y \end{pmatrix} \tag{2.2.1}$$

where $a_x$ and $a_y$ are the components along an orthogonal basis. The entries are complex and accordingly there are four independent parameters contained in (2.2.1). Since the entries are complex, they have magnitude and phase:

$$|a\rangle = \begin{pmatrix} |a_x|\mathrm{e}^{j\phi_x} \\ |a_y|\mathrm{e}^{j\phi_y} \end{pmatrix} = \mathrm{e}^{j\theta} \begin{pmatrix} |a_x| \\ |a_y|\mathrm{e}^{j\phi} \end{pmatrix} \tag{2.2.2}$$

where $\theta$ is a common phase and $\phi$ is the phase difference of the second row. In the following the explicit magnitude symbols $|\cdot|$ will be dropped and the intent of magnitude or complex number should be clear from the context. Bra vector $\langle a\,|$ is said to be the dual of $|a\rangle$ because they are not equal but they describe the same state:

$$|a\rangle \xleftrightarrow{\text{dual}} \langle a\,|$$

The bra vector $\langle a\,|$ corresponding to $|a\rangle$ is

$$\langle a\,| = \begin{pmatrix} a_x^* & a_y^* \end{pmatrix} \tag{2.2.3}$$

for every $|a\rangle$. The bra vector is the adjoint (†), or complex-conjugate transpose, of the corresponding ket vector:

$$\langle a\,| = (|a\rangle)^\dagger \tag{2.2.4}$$

Bra and ket vectors obey algebraic additive properties of identity, addition, commutation, and associativity. Identity and addition rules for kets are

$$\begin{aligned} &\text{identity} && |a\rangle + |0\rangle = |a\rangle \\ &\text{addition} && |a\rangle + |b\rangle = |\gamma\rangle \end{aligned}$$

where $|0\rangle$ is the null ket. Commutation and associativity are straightforward to prove using the matrix representation. A bra or ket vector can also be multiplied by a scalar quantity $c$:

$$c\,|a\rangle = |a\rangle\,c \tag{2.2.5}$$

Physically, the multiplication of a state vector by a scalar does not change the state and therefore the two commute. Operations that have no meaning are

the multiplication of multiple ket vectors or bra vectors. For example, $|b\rangle\,|a\rangle$ is meaningless.

Finally, it should be understood that state vectors $\langle a\,|$ and $|a\rangle$ are a more general representation than column and row vectors (2.2.1) and (2.2.3). A state vector is a coordinate-free abstraction that has the properties of length and direction; a row or column vector is a representation of a state vector given a choice of an underlying coordinate system.

### 2.2.2 Length and Inner Products

Bra and ket vectors have properties of length, phase, and pointing direction. The length of a real-valued vector is a scalar quantity and is determined by the dot product: $|\mathbf{a}|^2 = \mathbf{a}\cdot\mathbf{a}$. For complex-valued bra-ket vectors, the inner product is used to find length of a vector and is determined by multiplying its bra representation $\langle a\,|$ with its ket representation $|a\rangle$: $\|a\|^2 = \langle a\,|a\rangle$, where $\|\cdot\|$ is the norm of the vector.

More generally, one wants to measure the length of one vector as projected onto another. The inner product of two different vectors is the product of the bra form of one vector and the ket for of the other: $\langle b\,|a\rangle$. For real-valued vectors it is clear that $\mathbf{b}\cdot\mathbf{a} = \mathbf{a}\cdot\mathbf{b}$. However, for bra-ket vectors, having complex entries, the order of multiplication dictates the sign of the resulting phase. That is,

$$\langle b\,|a\rangle = |\langle b\,|a\rangle|\,\mathrm{e}^{j\gamma} \tag{2.2.6a}$$

$$\langle a\,|b\rangle = |\langle b\,|a\rangle|\,\mathrm{e}^{-j\gamma} \tag{2.2.6b}$$

The two inner products are related by the complex conjugate:

$$\langle b\,|a\rangle = (\langle a\,|b\rangle)^* \tag{2.2.7}$$

The inner product of a bra and ket is a complex-valued scalar. Based on (2.2.7) it is clear that the inner product of a vector onto itself yields a real number, and since the inner product is a measure of length, the real number is positive definite: $\langle a\,|a\rangle = \text{real number} \geq 0$. Only the null ket has length zero. Any finite ket has a length greater than zero. Throughout the body of the text, polarization vectors are taken to be unit vectors unless otherwise stated. A unit vector has a direction, phase, and unity length. Any vector can be converted to a unit vector by division by its norm:

$$|\tilde{a}\rangle = \frac{1}{\sqrt{\langle a\,|a\rangle}}\,|a\rangle \tag{2.2.8}$$

so that

$$\langle\tilde{a}\,|\tilde{a}\rangle = 1 \tag{2.2.9}$$

In the following the tilde over the vectors will be dropped.

Two vectors are defined as orthogonal to one another when the inner product vanishes:

$$\langle b \,|a\rangle = 0 \tag{2.2.10}$$

This is an essential inner product used regularly.

When two polarization vectors are resolved onto a common coordinate system,

$$\langle b \,|a\rangle = b_x^* a_x + b_y^* a_y \tag{2.2.11}$$

Finally, the inner product in matrix representation of a normalized vector is the sum of the component magnitudes squared:

$$\langle a \,|a\rangle = |a_x|^2 + |a_y|^2 = 1 \tag{2.2.12}$$

### 2.2.3 Projectors and Outer Products

The inner product measures the length of a vector or the projection of one vector onto another. The result is a complex scalar quantity. In contrast, the outer product retains a vector nature while also producing length by projection. There are two outer product types to study: the projector, having the form $|p\rangle\langle p|$; and the outer product $|p\rangle\langle q|$. The form $|p\rangle\langle q|$ is called a dyadic pair because the vector pair has neither a dot nor cross product between them. In quantum mechanics the projector $|p\rangle\langle p|$ is called the density operator for the state.

Consider a projector that operates on ket $|a\rangle$:

$$|p\rangle\langle p \,|a\rangle = |p\rangle \left(\langle p \,|a\rangle\right) = c \,|p\rangle \tag{2.2.13}$$

The quantity $c = \langle p \,|a\rangle$ is just a complex scalar and commutes with the ket. Operating on $|a\rangle$ the projector measures the length of $|a\rangle$ on $|p\rangle$ and produces a new vector $|p\rangle$.

The effect of the projector is to point along the $|p\rangle$ direction where the length of $|p\rangle$ is scaled by $\langle p \,|a\rangle$. Projectors work equally well on bras, e.g.

$$\langle a \,|p\rangle \,\langle p \,| = c^* \langle p \,| \tag{2.2.14}$$

so in fact it should be clear that

$$\langle a \,|p\rangle \,\langle p \,| = \left(|p\rangle\langle p \,|a\rangle\right)^\dagger \tag{2.2.15}$$

The adjoint operator connects the bra and ket forms.

The behavior of the outer product $|p\rangle\langle q|$ is similar to the projector but for the fact that the projection vector and resultant pointing direction differ. The resultant pointing direction depends whether the outer product operates on a ket or a bra. Acting on a ket, the outer product yields

$$|p\rangle\langle q \,|a\rangle = \left(\langle q \,|a\rangle\right) |p\rangle \tag{2.2.16}$$

whereas acting on a bra of the same vector, the outer product yields

$$\langle a\,|p\rangle\,\langle q\,| = (\langle a\,|p\rangle)\,\langle q\,| \tag{2.2.17}$$

The resultant pointing direction and projected length depends on whether the outer product operates on a bra or ket vector.

In the study of polarization, the outer product is a $2 \times 2$ matrix with complex entries:

$$|b\rangle\langle a| = \begin{pmatrix} b_x a_x^* & b_x a_y^* \\ b_y a_x^* & b_y a_y^* \end{pmatrix} \tag{2.2.18}$$

The determinant is

$$\det(|b\rangle\langle a|) = 0 \tag{2.2.19}$$

and therefore the projector is non-invertible. The determinant of an outer product of any dimension is likewise zero. That means the action of $|b\rangle\langle a|$ on a ket is irreversible, which is reasonable because the original direction of the ket is lost. So, while all outer products are operators not all operators are outer products. Operators that are linear combinations of projectors are reversible under the right construction.

In summary, the outer product follows these rules:

| | |
|---|---|
| equivalence | $(\|b\rangle\langle a\|)^\dagger = \|a\rangle\langle b\|$ |
| associative | $(\|b\rangle\langle a\|)\,\|\gamma\rangle = \langle b\,|\,(\langle a\,|\gamma\rangle)$ |
| trace | $\mathrm{Tr}\,(\|b\rangle\langle a\|) = \langle a\,|b\rangle$ |
| irreversible | $\det(\|b\rangle\langle a\|) = 0$ |

where Tr stands for the trace operation. The trace connects the outer product to the inner product.

### 2.2.4 Orthonormal Basis

An orthonormal basis is a complete set of orthogonal unit-length axes on which any vector in the space can be resolved. Consider a vector space with $N$ dimensions and orthogonal unit vectors $(|a_1\rangle, |a_2\rangle, \ldots, |a_N\rangle)$. The orthogonality requires

$$\langle a_m\,|a_n\rangle = \delta_{m,n} \tag{2.2.20}$$

where $\delta m, n$ is the Kronecker delta function. Only a vector projected onto itself yields a non-vanishing inner product. When the set is complete, the outer products are closed, where closure is defined as

$$\sum_n |a_n\rangle\langle a_n| = I \tag{2.2.21}$$

When a basis set, or group, is closed, any operation to a member of the group results in another member within the group. Together, (2.2.20–2.2.21) are the two conditions that define an orthonormal basis.

Given an orthonormal basis, any arbitrary vector can be resolved onto the basis using (2.2.21). An arbitrary ket $|s\rangle$ is resolved as

$$|s\rangle = \left( \sum_n |a_n\rangle\langle a_n| \right) |s\rangle = \sum_n c_n |a_n\rangle \tag{2.2.22}$$

where the complex coefficients are given by $c_n = \langle a_n | s\rangle$. The inner product $\langle s | s\rangle$ is the sum of the absolute-value squares of the coefficients $c_n$:

$$\langle s | s\rangle = \sum_{a'} |c_{a'}|^2 \tag{2.2.23}$$

When $|s\rangle$ is normalized $\sum_{a'} |c_{a'}|^2 = 1$.

## 2.3 General Vector Transformations

Interaction between a physical system and external influences can change the state of a system. Left unperturbed, a state persists indefinitely. Operators embody the action of external influences and are distinct from the state of the system itself. The bra and ket vectors of the preceding section are two equivalent spaces that describe the same state space. Operators also have two distinct and equivalent spaces that describe the same state transformation. While there is no special notation to represent a "ket" operator or a "bra" operator, equivalence between operator spaces is maintained under

$$X |a\rangle \xleftrightarrow{\text{dual}} \langle a | X^\dagger \tag{2.3.1}$$

$X^\dagger$ is said to be the adjoint operator of $X$. Care should be taken because the action of $X |a\rangle$ is not the same as $\langle a | X$; these two results are different.

### 2.3.1 Operator Relations

Operators always act on kets from the left and bras from the right, e.g. $X |a\rangle$ or $\langle a | X$. The expressions $|a\rangle X$ and $X\langle a |$ are undefined. An operator multiplying a ket produces a new ket, and an operator multiplying a bra produces a new bra. In general, an operator changes the state of the system,

$$X |a\rangle = c |b\rangle \tag{2.3.2}$$

where $c$ is a scaling factor induced solely by $X$. Operators are said to be equal if

$$X |a\rangle = Y |a\rangle \Rightarrow X = Y \tag{2.3.3}$$

Operators obey the following arithmetic properties of addition:

| | |
|---|---|
| commutative | $X + Y = Y + X$ |
| associative | $X + (Y + Z) = (X + Y) + Z$ |
| distributive | $X\left(\lvert a\rangle + \lvert b\rangle\right) = X\lvert a\rangle + X\lvert b\rangle$ |

Operators in general do not commute under multiplication. That is

$$XY \neq YX \tag{2.3.4}$$

In matrix form, only when $X$ and $Y$ are diagonal matrices does $XY = YX$. Other multiplicative properties are

| | |
|---|---|
| identity | $I\lvert a\rangle = \lvert a\rangle$ |
| associative | $X(YZ) = (XY)Z$ |
| distributive | $X\left(Y\lvert a\rangle\right) = XY\lvert a\rangle$ |

All of the above arithmetic properties apply equally well to bra vectors.

The effect $X$ has on state $\vec{a}$ is measured by

$$\text{expectation value of } X \text{ on } \vec{a} = \frac{\langle a\lvert X\rvert a\rangle}{\langle a\lvert a\rangle} \tag{2.3.5}$$

In general, an inner product that encloses an operator gives a complex number:

$$\langle b\rvert\left(X\lvert a\rangle\right) = \langle b\lvert X\rvert a\rangle = \text{complex number} \tag{2.3.6}$$

Consider dual constructions, first where $X\lvert a\rangle$ is left-multiplied by $\langle b\rvert$, and second where the dual $\langle a\rvert X^\dagger$ is right-multiplied by $\lvert b\rangle$:

$$\langle b\lvert X\rvert a\rangle = \left(\langle a\rvert X^\dagger\rvert b\rangle\right)^* \tag{2.3.7}$$

These two cases are duals of one another and are therefore complex conjugates.

In the study of polarization, operators are represented as $2 \times 2$ complex-valued Jones matrices:

$$X = \begin{pmatrix} a\,e^{j\alpha} & b\,e^{j\beta} \\ c\,e^{j\gamma} & d\,e^{j\eta} \end{pmatrix} \tag{2.3.8}$$

There are eight independent variables contained in the operator. If $\det X \neq 0$, then $X$ is invertible and the action of $X$ can be undone.

The properties of operators are summarized as follows:

| | |
|---|---|
| operator duality | $X\lvert a\rangle \xleftrightarrow{\text{dual}} \langle a\rvert X^\dagger$ |
| change of state | $X\lvert a\rangle = c\lvert b\rangle$ |
| | $\langle a\rvert X^\dagger = c^*\langle b\rvert$ |
| inner product with operator | $\langle b\rvert X\rvert a\rangle = \text{complex number}$ |
| conjugate relation | $\langle b\rvert X\rvert a\rangle = \langle a\rvert X^\dagger\rvert b\rangle^*$ |
| conjugate transpose | $(XY)^\dagger = Y^\dagger X^\dagger$ |

Just as an arbitrary ket can be resolved onto an orthonormal basis, an arbitrary operator $X$ can be resolved onto a set of projection operators formed on the orthonormal basis. Applying the closure relation (2.2.21) yields

$$\begin{aligned} X &= \left(\sum_m |a_m\rangle\langle a_m|\right) X \left(\sum_n |a_n\rangle\langle a_n|\right) \\ &= \sum_n \sum_m |a_m\rangle \langle a_m |X| a_n\rangle \langle a_n| \end{aligned} \tag{2.3.9}$$

The indexing symmetry of (2.3.9) looks like a matrix with $\langle a_m |X| a_n\rangle$ as the $(m, n)$ entry. For polarization, the matrix is $2 \times 2$ and looks like

$$\sum_n \sum_m |a_m\rangle \langle a_m |X| a_n\rangle \langle a_n| \mapsto \begin{pmatrix} \langle a_1 |X| a_1\rangle & \langle a_1 |X| a_2\rangle \\ \langle a_2 |X| a_1\rangle & \langle a_2 |X| a_2\rangle \end{pmatrix} \tag{2.3.10}$$

The resolved form of $X$ in (2.3.10) will become particularly simple in discussion of Hermitian and unitary matrices.

## 2.4 Eigenstates, Hermitian and Unitary Operators

Many physical systems exhibit particular states that are not transformed by interaction with the system. These invariant states are called eigenstates of the system. In spin-vector calculus, operators embody the influence of a phenomena. The eigenvectors of an operator are the eigenstates of the system. When an operator $X$ acts on its own eigenstate $\vec{a}$,

$$X |a_1\rangle = a_1 |a_1\rangle \tag{2.4.1a}$$

$$\langle a_1 | X^\dagger = a_1^* \langle a_1 | \tag{2.4.1b}$$

the state of the system is unaltered but for a scaling factor $a_1$. The scale factor is the eigenvalue of $X$ associated with eigenstate $\vec{a}_1$. Each eigenvector has an associated eigenvalue, and a well-conditioned matrix has as many eigenvectors as rows in the matrix or, equivalently, dimensions in the state space.

The eigenvectors of Hermitian and unitary operators are orthogonal when the associated eigenvalues are distinct. The eigenvalues of a Hermitian operator are real-valued scalars, and the eigenvalues of a unitary operator are complex exponential scalars. A Hermitian or unitary operator $X$ having $N$ eigenkets $(|a_1\rangle, |a_2\rangle, \ldots, |a_N\rangle)$ and associated eigenvalues $(a_1, a_2, \ldots, a_N)$ produces the series of inner products

$$\langle a_m \left| X^\dagger X \right| a_n\rangle = |a_m|^2 \, \delta_{m,n} \tag{2.4.2}$$

The operator $X^\dagger X$ scales each axis by a different amount, but does not rotate nor create reflection of the original basis. The eigenvalues of operator $X$ are related to the determinant and trace by

$$\det(X) = a_1\, a_2\, \cdots a_N \tag{2.4.3a}$$

$$\mathrm{Tr}(X) = a_1 + a_2 + \cdots + a_N \tag{2.4.3b}$$

Since the eigenvalues of a Hermitian matrix are real, its determinant and trace are real.

### 2.4.1 Hermitian Operators

The defining property of a Hermitian operator is

$$H^\dagger = H \tag{2.4.4}$$

The associated Hermitian matrix in polarization studies has only four independent variables: three amplitudes and one phase. This contrasts with the general Jones matrix (2.3.8) which has eight.

The eigenvectors of $H$ form a complete orthonormal basis and the eigenvalues are real. That the eigenvalues are real is proved from the following difference:

$$\begin{aligned} \langle a_n \left| \left(H^\dagger - H\right) \right| a_m \rangle &= ({a_n}^* - a_m)\, \langle a_n \, | a_m \rangle \\ &= 0 \end{aligned} \tag{2.4.5}$$

Non-trivial solutions are found when neither vector is null. The eigenvectors may be the same or different. Consider first when the eigenvectors are the same. Since $\langle a_n \, | a_n \rangle \neq 0$, $(a_n^* - a_n) = 0$ and the eigenvalue is real. Consider when the eigenvectors are different. Unless $a_m = a_n$, in which case the eigenvectors are not linearly independent, it must be the case that $\langle a_n \, | a_m \rangle = 0$. All eigenvalues are therefore real. Hermitian operators $H$ scale its own basis set:

$$\langle a_m \left| H^\dagger H \right| a_n \rangle = a_m^2\, \delta_{m,n} \tag{2.4.6}$$

When $\det(H) \neq 0$, $H$ is invertible and the action of $H$ on the state of a system is reversible.

The expansion of $H$ onto its own basis generates a diagonal eigenvalue matrix. Under construction (2.3.9) the expansion yields

$$\begin{aligned} H &= \sum_n \sum_m |a_m\rangle \, \langle a_m \, |H| \, a_n \rangle \langle a_m \, | \\ &= \sum_m a_m \, |a_m\rangle\langle a_m| \end{aligned} \tag{2.4.7}$$

where $\langle a_m \, |H| \, a_n \rangle = a_n \delta_{m,n}$. The orthonormal expansion is written in matrix form as $H = S\Lambda S^{-1}$, where $S$ is a square matrix whose columns are the eigenvectors of $H$ and $\Lambda$ is a diagonal matrix whose entries are the associated eigenvalues. Schematically,

$$S = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_N \\ | & | & & | \end{pmatrix}, \quad \text{and} \quad \Lambda = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_N \end{pmatrix} \tag{2.4.8}$$

where $|a_n\rangle = v_n^T$.

### 2.4.2 Unitary Operators

The defining property of a unitary operator is

$$T^\dagger T = I \tag{2.4.9}$$

Acting on its orthogonal eigenvectors $|a_n\rangle$, the unitary operator preserves the unity basis length:

$$\langle a_m \left| T^\dagger T \right| a_n \rangle = \delta_{m,n} \tag{2.4.10}$$

Taking the determinant of both sides of (2.4.9) gives $\det(T^\dagger T) = 1$. Since the determinant of a product is the product of the determinants and the adjoint operator preserves the norm, the determinant of $T$ must be

$$\det(T) = \mathrm{e}^{j\theta} \tag{2.4.11}$$

Since the the determinant is the product of eigenvalues, the eigenvalues of $T$ must themselves be complex exponentials and, accounting for (2.4.10), they must have unity magnitude. Therefore $T$ acting on an eigenvector yields

$$T\,|a_n\rangle = \mathrm{e}^{-j\alpha_n}\,|a_n\rangle \tag{2.4.12}$$

The eigenvalues of $T$ lie on the unit circle in the complex plane.

A special form of $T$ exists where the determinant is unity. This special form is denoted $U$ and is characterized by $\det(U) = +1$. To transform from $T$ to $U$, the common phase factor $\beta = \exp(j\theta/N)$ must be extracted from each eigenvalue of $T$, where $N$ is the dimensionality of the operator. The $T$ and $U$ forms are thereby related:

$$T = \mathrm{e}^{j\beta} U \tag{2.4.13}$$

It should be noted that when $\det(U) = -1$, a reflection is present along an odd number of axes in the basis set of $U$.

The eigenvalue equation for $U$ is

$$U\,|a_n\rangle = \mathrm{e}^{-j\phi_n}\,|a_n\rangle \tag{2.4.14}$$

$U$ expands on its own basis set in the same way $H$ expands (2.4.7):

$$U = \sum_m \mathrm{e}^{-j\phi_m}\,|a_m\rangle\langle a_m| \tag{2.4.15}$$

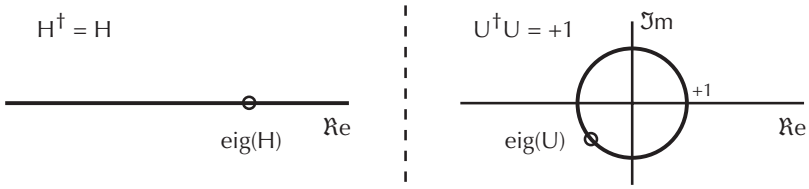


**Fig. 2.1.** Eigenvalue loci of $H$ and $U$. Left: eigenvalues of $H$ lie on the real number line. Right: eigenvalues of $U$ lie on the unit circle in the complex plane.

This orthonormal expansion has the matrix analogue of $U = S\exp(-j\Lambda)S^{-1}$, where the diagonal matrix is

$$\exp(-j\Lambda) = \begin{pmatrix} \mathrm{e}^{-j\phi_1} & & & \\ & \mathrm{e}^{-j\phi_2} & & \\ & & \ddots & \\ & & & \mathrm{e}^{-j\phi_N} \end{pmatrix} \tag{2.4.16}$$

and $S$ is the same form as (2.4.8).

### 2.4.3 Connection between Hermitian and Unitary Matrices

The connection between Hermitian and unitary operators is quite intimate. Figure 2.1 illustrates the eigenvalue domains for $H$ and $U$. The eigenvalues of $H$ lie on the real number line, while those of $U$ lie on the unit circle in the complex plane. Multiplying the eigenvalues of $H$ by $-j$ and taking the exponential, one can construct the eigenvalues of $U$. Note that the eigenvalues of $U$ are cyclic, so only the real number line modulo $2\pi$ is significant.

Based on the operator expansions of the preceding sections, one has

$$H = S\Lambda_H S^{-1} \tag{2.4.17a}$$

$$U = R\,\mathrm{e}^{-j\Lambda_U}R^{-1} \tag{2.4.17b}$$

Since in general $\exp(S\Lambda S^{-1}) = S\exp(\Lambda)S^{-1}$, the $H$ and $U$ operators may be connected as

$$U = \mathrm{e}^{-jH} \quad \Longrightarrow \quad S\mathrm{e}^{-j\Lambda_U}S^{-1} = S\mathrm{e}^{-j\Lambda_H}S^{-1} \tag{2.4.18}$$

For every Hermitian operator $H$ there is an associated unitary operator $U$ that shares the same basis set and has eigenvalues related through the complex exponential.

### 2.4.4 Similarity Transforms

Frequently one has Hermitian operator $H$ and orthonormal basis $|p_n\rangle$ that are not aligned. That is, the eigenvectors $|a_n\rangle$ of $H$ are not parallel to vectors $|p_n\rangle$. Expansion of $H$ into $|p_n\rangle$ using the expansion expression (2.3.10)

generates a matrix that is not diagonal. However, the expansion matrix can be diagonalized by rotating basis $|p_n\rangle$ into $|a_n\rangle$. The unitary matrix does this operation. Taking advantage of $U^\dagger U = 1$, one can write

$$\begin{aligned} \langle p\,|H|\,p\rangle &= \langle p\,\big|U^\dagger U H U^\dagger U\big|\,p\rangle \\ &= \langle a\,\big|UHU^\dagger\big|\,a\rangle \\ &= \langle a\,|H_T|\,a\rangle \end{aligned} \tag{2.4.19}$$

Since (2.4.19) holds for any choice of initial basis $|p_n\rangle$, the operators

$$H_T = UHU^\dagger \tag{2.4.20}$$

must be equal. Equation (2.4.20) is known as a similarity transform. Both the determinant and trace of $H$ are independent of basis; that is

$$\begin{aligned} \det(H_T) &= \det(U)\det(H)\det(U^\dagger) \\ &= \det(H) \end{aligned} \tag{2.4.21}$$

and

$$\mathrm{Tr}(H_T) = \mathrm{Tr}(UHU^\dagger) \tag{2.4.22}$$

The trace is always preserved under a similarity transform.

### 2.4.5 Construction of General Unitary Matrix

The characteristic property $T^\dagger T = 1$ of unitary matrices restricts the eight independent variables generally available for a polarization operator (2.3.8). Derivation of the restrictions generates a general form of the unitary matrix.

First consider $U$, where $\det(U) = +1$ and . Substitution of $X$ (2.3.8) for $U$ in $U^\dagger U = 1$ generates the following requirements:

$$\begin{gathered} |a|^2 + |b|^2 = 1, \quad |a|^2 = |d|^2, \quad |b|^2 = |c|^2, \\ ac\,\mathrm{e}^{-j(\alpha-\gamma)} + bd\,\mathrm{e}^{-j(\beta-\eta)} = 0 \end{gathered} \tag{2.4.23}$$

The determinant requirement generates

$$ad\,\mathrm{e}^{j(\alpha+\eta)} - bc\,\mathrm{e}^{j(\beta+\gamma)} = 1 \tag{2.4.24}$$

The amplitude restrictions in (2.4.23) are satisfied by $a = \cos\kappa$ and $b = \sin\kappa$. There remains, however, a sign degeneracy in that $c = \pm b$ and $d = \pm a$. This degeneracy is insignificant in that any choice flows through the restriction criteria and produces the same matrix form.

Combination of the last equation in (2.4.23) and (2.4.24) generates two restrictions on the phase:

$$\begin{aligned} e^{j(\gamma+\beta)} &= -e^{j(\alpha+\eta)} \\ e^{j\alpha} &= e^{-j\eta} \\ e^{j\gamma} &= -e^{-j\beta} \end{aligned} \tag{2.4.25}$$

There are only two independent phases. Combining all of the above restrictions, the general matrix form of $U$ is written

$$U = \begin{pmatrix} e^{j\alpha}\cos\kappa & -e^{j\beta}\sin\kappa \\ e^{-j\beta}\sin\kappa & e^{-j\alpha}\cos\kappa \end{pmatrix} \tag{2.4.26}$$

There are three independent variables in $U$: one amplitude and two phases. The fourth independent variable has been suppressed because of the arbitrary selection $\det(U) = +1$. The unitary matrix $T$ includes the common phase:

$$T = e^{j\phi}\begin{pmatrix} e^{j\alpha}\cos\kappa & -e^{j\beta}\sin\kappa \\ e^{-j\beta}\sin\kappa & e^{-j\alpha}\cos\kappa \end{pmatrix} \tag{2.4.27}$$

where there are now four independent variables: one amplitude and three phases.

The Cayley-Klein form of $U$, using complex entries $a$ and $b$, is

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \tag{2.4.28}$$

The inverse of the unitary matrix is $U^{-1}(a,b) = U(a^*,-b)$.

### 2.4.6 Group Properties of SU(2)

For polarization studies, unitary operators are $2 \times 2$ square matrices with complex entries. The group defined by multiplication operations of $2 \times 2$ unitary matrices is called U(2), and the subgroup of unitary matrices where $\det(U) = +1$ is called SU(2), "S" for special. The group properties for multiplication are

| | |
|---|---|
| Identity | $UI = U$ |
| Closure | $U_1U_2 = U_3$ |
| Inverse | $U^{-1}U = I$ |
| Associativity | $(U_1U_2)U_3 = U_1(U_2U_3)$ |

where in all cases $U_{1,2,3} \in \mathrm{SU}(2)$. SU(2) is closed under these four operations.

## 2.5 Vectors Cast in Jones and Stokes Spaces

Thus far, state spaces and operators have been presented without restriction on their dimensionality. The properties of these vectors and matrices have been studied in general with passing observations about polarization-specific points of interest. At this stage the scope of presentation will concentrate on the study of polarization so that a formal connection between Jones and Stokes space can be established. The tools developed in the preceding sections are essential to make the bilateral connections that follow.

Recall from (1.4.5) on page 13 that a polarization vector is written as

$$|s\rangle = E_o \, \mathrm{e}^{j\theta} \begin{pmatrix} \cos\chi \\ \sin\chi \mathrm{e}^{j\phi} \end{pmatrix} \tag{2.5.1}$$

where $E_o$ is real. There are two polar angles in (2.5.1): $\chi$ and $\phi$. The common phase $\exp(j\theta)$ is lost on conversion to Stokes space.

There are seven measurements necessary to determine the polarization ellipse uniquely. The first measurement is for the overall intensity and the remaining measurements project the ellipse onto six different reference axes. The formal construction of a projection matrix is necessary at this point.

Consider points along two orthogonal axes and their projection onto a line $L$ inclined by angle $\theta$ that passes through the origin. As illustrated in Fig. 2.2, the coordinate $(1, 0)$ is projected to point $a$ on line $L$. The coordinates of $a$ as measured along the two orthogonal axes are $(\cos^2\theta, \sin\theta\cos\theta)$. After a similar analysis for the coordinate $(0, 1)$, one can construct the projection matrix $\mathcal{P}$:

$$\mathcal{P} = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix} \tag{2.5.2}$$

It is clear that $\det(\mathcal{P}) = 0$; $\mathcal{P}$ is non-invertable and its action is irreversible. There is loss of information after projection. Moreover, $\mathcal{P}^2 = \mathcal{P}$, so once the projection is taken, subsequent projections along the same line $L$ do not change the result.

### 2.5.1 Complete Measurement of the Polarization Ellipse

There are seven measurements necessary for complete determination of the polarization ellipse. The first measurement is one of total intensity, the remaining six measurements are projections. The projections are defined in pairs and the difference values are associated with the Stokes coordinates. The result of an intensity of the polarization ellipse is the inner product

$$\langle s\,|s\rangle = \langle s\,| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} |s\rangle \tag{2.5.3}$$

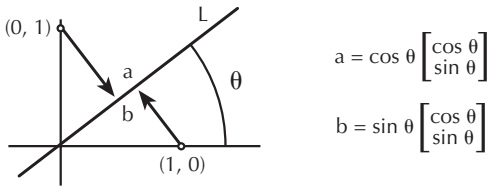


**Fig. 2.2.** Projection of unit coordinates $(1,0)$ and $(0,1)$ onto line $L$, which is inclined by angle $\theta$ and passes through the origin. The projected coordinates are tabulated on the right. A second projection of $a$ and $b$ onto $L$ does not change the coordinates of $a$ and $b$. The projection operator is non-invertable.

Without loss of generality, $\langle s\,|s\rangle = 1$ in the following.

The first projection pair is $\theta = 0$ and $\theta = \pi/2$. The projection measure comes from the inner products

$$P_0 = \langle s\,| \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} |s\rangle, \quad \text{and} \quad P_{\pi/2} = \langle s\,| \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} |s\rangle \tag{2.5.4}$$

The Stokes parameter $s_1$ is defined as

$$s_1 = P_0 - P_{\pi/2} \tag{2.5.5}$$

Substitution of (2.5.4) into (2.5.5) makes

$$s_1 = \langle s\,| \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} |s\rangle \tag{2.5.6}$$

The second projection pair is $\theta = \pm\pi/4$. These projections produce

$$P_{+\pi/4} = \tfrac{1}{2}\,\langle s\,| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} |s\rangle, \quad \text{and} \quad P_{-\pi/4} = \tfrac{1}{2}\,\langle s\,| \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} |s\rangle \tag{2.5.7}$$

The Stokes parameter $s_2$ is defined as

$$s_2 = P_{+\pi/4} - P_{-\pi/4} \tag{2.5.8}$$

which makes

$$s_2 = \langle s\,| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} |s\rangle \tag{2.5.9}$$

The last projection requires the measurement of the ellipse circularity. By convention, right-hand circular polarization rotates in the counter-clockwise (ccw) direction when observed along the $-\hat{z}$ direction (looking into the light). The right-hand circular polarization vector is

$$|s\rangle_R = \begin{pmatrix} 1 \\ j \end{pmatrix} \tag{2.5.10}$$

The ccw vector needs mapping to the $\theta = 0$ axis; a unitary transform does the rotation. The right- and left-hand projections are calculated via

$$P_R = \langle s\,|\,U^\dagger \mathcal{P}_0 U\,|\,s\rangle \tag{2.5.11a}$$

$$P_L = \langle s\,|\,U^\dagger \mathcal{P}_{\pi/2} U\,|\,s\rangle \tag{2.5.11b}$$

The unitary matrix

$$U = \frac{1}{2}\begin{pmatrix} 1 & -j \\ -j & 1 \end{pmatrix} \tag{2.5.12}$$

maps right-hand circular polarization to the $\theta = 0$ axis:

$$\frac{1}{2}\begin{pmatrix} 1 & -j \\ -j & 1 \end{pmatrix}\begin{pmatrix} 1 \\ j \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.5.13}$$

Substituting (2.5.2) and (2.5.12) into (2.5.11) produces

$$P_R = \tfrac{1}{4}\,\langle s\,|\begin{pmatrix} 1 & -j \\ j & 1 \end{pmatrix}|\,s\rangle\,, \quad \text{and} \quad P_L = \tfrac{1}{4}\,\langle s\,|\begin{pmatrix} 1 & j \\ -j & 1 \end{pmatrix}|\,s\rangle \tag{2.5.14}$$

The Stokes parameter $s_3$ is defined as

$$s_3 = P_R - P_L \tag{2.5.15}$$

which makes

$$s_3 = \langle s\,|\begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}|\,s\rangle \tag{2.5.16}$$

From these seven measurements one can transform from a ket in Jones space to three Stokes coordinates that lie on the unit sphere:

$$|s\rangle \Longrightarrow \hat{s} \tag{2.5.17}$$

where the vector $\hat{s}$ is a column vector defined by $\hat{s} = (s_1, s_2, s_3)^T$.

This completes the measurement of the polarization ellipse. From these measurements the polarimetric angles $\chi$ and $\phi$ are uniquely determined. These measurements combined with the definition of the Stokes parameters generate the Pauli spin matrices. This is the topic of the next section.

### 2.5.2 Pauli Spin Matrices

The Pauli spin matrices connect Jones to Stokes spaces through the projection measurements of the preceding section. The identity Pauli matrix is

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.5.18}$$

The Pauli spin matrices are[1]

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} \tag{2.5.19}$$

The spin matrices are both Hermitian and unitary:

$$\sigma_k{}^\dagger = \sigma_k \quad \text{and} \quad \sigma_k{}^\dagger \sigma_k = I \tag{2.5.20}$$

The determinants of the spin matrices are $-1$ and the traces zero:

$$\det(\sigma_k) = -1 \quad \text{and} \quad \text{Tr}(\sigma_k) = 0 \tag{2.5.21}$$

A spin matrix multiplied by itself yields

$$\sigma_k \sigma_k = I \tag{2.5.22}$$

and multiplied by other matrices gives

$$\sigma_i \sigma_j = -\sigma_j \sigma_i = j\sigma_k \tag{2.5.23}$$

where the indices of the multiplication table $(i, j, k)$ are cyclic permutations of $(1, 2, 3)$.

Each Stokes coordinate of a polarization state $|s\rangle$ is calculated by inserting the associated Pauli matrix into the inner product $\langle s\,|\cdot|\,s\rangle$. The individual Stokes coordinates are

$$s_k = \langle s\,|\sigma_k|\,s\rangle \tag{2.5.24}$$

This is shorthand for the projection-difference measurements of (2.5.6, 2.5.9, 2.5.16). Since the spin matrices are Hermitian, the Stokes coordinates $s_k$ are real, signed quantities. Moreover, since $\det(\sigma_k) = -1$ and the Jones vector $|s\rangle$ is assumed to be normalized, $s_k$ is bounded by $-1 \le s_k \le +1$. The proof that the norm of $\hat{s}$ is unity, $|\hat{s}| = 1$, is shown below.

### 2.5.3 The Pauli Spin Vector and the Bilateral Connection Between Jones and Stokes Vectors

The Pauli spin vector condenses further the notation of (2.5.24). The spin vector is defined as

$$\vec{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} \tag{2.5.25}$$

[1] In physics texts the $z$ direction is denoted by the $\sigma_1$ spin matrix while here it is denoted by $\sigma_3$. Historically, the Pauli spin matrices describe electron spin, which is either up or down in the "$z$" direction. In polarization optics, one usually thinks of a horizontal polarization state aligned to the "$x$" axis.

where $\vec{\sigma}$ is a vector of matrices. The vector of Stokes coordinates $\hat{s}$ is derived from the Jones vector $|s\rangle$ using the spin vector:

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} \langle s|\sigma_1|s\rangle \\ \langle s|\sigma_2|s\rangle \\ \langle s|\sigma_3|s\rangle \end{pmatrix} \tag{2.5.26}$$

More concisely,

$$\hat{s} = \langle s|\vec{\sigma}|s\rangle \tag{2.5.27}$$

This is the most compact way to map Jones vectors to Stokes vectors.

The reciprocal connection is made through an eigenvalue equation whose parameters are the Stokes vector $\hat{s}$ and the spin vector. First, observe that the spin vector behaves both as a $3\times1$ vector and as a $2\times2$ matrix, depending on the context. Above shows the spin vector acting as a $3\times1$ vector. Alternatively, the dot product of $\hat{s}$ with the spin vector yields

$$\begin{aligned} \hat{s}\cdot\vec{\sigma} &= s_1\sigma_1 + s_2\sigma_2 + s_3\sigma_3 \\ &= \begin{pmatrix} s_1 & s_2 - js_3 \\ s_2 + js_3 & -s_1 \end{pmatrix} \end{aligned} \tag{2.5.28}$$

$\hat{s}\cdot\vec{\sigma}$ in this case is a $2\times2$ Jones matrix and, since the coefficients $s_k$ are real, $\hat{s}\cdot\vec{\sigma}$ is Hermitian: $(\hat{s}\cdot\vec{\sigma})^\dagger = (\hat{s}\cdot\vec{\sigma})$.

Next, recall from §2.2.3 that the trace operation connects the projector with its inner product: $\mathrm{Tr}(|s\rangle\langle s|) = \langle s|s\rangle$. Since the trace of each Pauli matrix is zero it is also true that $\mathrm{Tr}\,(\hat{s}\cdot\vec{\sigma}) = 0$. For a normalized state vector such that $\langle s|s\rangle = 1$, one can construct the projector for ket $|s\rangle$ in terms of the spin vector:

$$|s\rangle\langle s| = \frac{1}{2}\left(I + \hat{s}\cdot\vec{\sigma}\right) \tag{2.5.29}$$

Subsequent multiplication on the right by $|s\rangle$ generates the eigenvalue equation

$$\hat{s}\cdot\vec{\sigma}\,|s\rangle = |s\rangle \tag{2.5.30}$$

This is the most compact way to map Stokes vectors to Jones vectors. The eigenvector of $\hat{s}\cdot\vec{\sigma}$ associated with eigenvalue $+1$ generates the Jones vector $|s\rangle$ from Stokes vector $\hat{s}$.

### 2.5.4 Spin-Vector Identities

Vector operations that include spin-vectors do not yield to the same intuition one is accustomed to with "normal" vectors. For example, while one is quite familiar with $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{a}) = 0$, since $\mathbf{a}$ is orthogonal to $\mathbf{b}\times\mathbf{a}$, the spin-vector analogue produces $\vec{\sigma}\cdot(\vec{a}\times\vec{\sigma}) = -2j(\vec{a}\cdot\vec{\sigma})$. The difference comes from

the cyclic multiplication table for spin-vectors (2.5.22–2.5.23), where the sign of a product is determined by the order in which the spin-vectors appear.

The purpose of the following identity tabulation is to provide reductions in the order $k$ of $(\vec{\sigma})^k$. For the following identities, $\vec{a}$ and $\vec{b}$ are real-valued $3\times 1$ vectors and $\vec{\sigma}$ is the spin vector. Real vectors $\vec{a}$ and $\vec{b}$ are not interchangeable with the spin vector $\vec{\sigma}$.

Identities of order $(\vec{\sigma})^0$ and $(\vec{\sigma})$:

$$\vec{a}\cdot\vec{a} = a^2 \tag{2.5.31}$$

$$\vec{a}\cdot\vec{\sigma} = \vec{\sigma}\cdot\vec{a} \tag{2.5.32}$$

$$\vec{a}(\vec{a}\cdot\vec{\sigma}) = (\vec{a}\cdot\vec{\sigma})\vec{a} \tag{2.5.33}$$

Identities of order $(\vec{\sigma})^2$:

$$\vec{\sigma}\cdot\vec{\sigma} = 3I \tag{2.5.34}$$

$$\vec{\sigma}(\vec{a}\cdot\vec{\sigma}) = \vec{a}I + j\vec{a}\times\vec{\sigma} \tag{2.5.35}$$

$$(\vec{a}\cdot\vec{\sigma})\vec{\sigma} = \vec{a}I - j\vec{a}\times\vec{\sigma} \tag{2.5.36}$$

$$(\vec{a}\cdot\vec{\sigma})(\vec{a}\cdot\vec{\sigma}) = a^2\, I \tag{2.5.37}$$

$$(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma}) = (\vec{a}\cdot\vec{b})I + (j\vec{a}\times\vec{b})\cdot\vec{\sigma} \tag{2.5.38}$$

$$[(\vec{a}\cdot\vec{\sigma}),\vec{\sigma}] = -2j\vec{a}\times\vec{\sigma} \tag{2.5.39}$$

$$\{(\vec{a}\cdot\vec{\sigma}),\vec{\sigma}\} = 2\vec{a}\, I \tag{2.5.40}$$

$$\Big[(\vec{a}\cdot\vec{\sigma}),(\vec{b}\cdot\vec{\sigma})\Big] = 2(j\vec{a}\times\vec{b})\cdot\vec{\sigma} \tag{2.5.41}$$

$$\Big\{(\vec{a}\cdot\vec{\sigma}),(\vec{b}\cdot\vec{\sigma})\Big\} = 2(\vec{a}\cdot\vec{b})\, I \tag{2.5.42}$$

$$\vec{\sigma}\cdot(j\vec{a}\times\vec{\sigma}) = 2(\vec{a}\cdot\vec{\sigma}) \tag{2.5.43}$$

$$(j\vec{a}\times\vec{\sigma})\cdot\vec{\sigma} = -2(\vec{a}\cdot\vec{\sigma}) \tag{2.5.44}$$

$$(j\vec{a}\times\vec{\sigma})(\vec{a}\cdot\vec{\sigma}) = a^2\vec{\sigma} - \vec{a}(\vec{a}\cdot\vec{\sigma}) \tag{2.5.45}$$

$$(\vec{a}\cdot\vec{\sigma})(j\vec{a}\times\vec{\sigma}) = \vec{a}(\vec{a}\cdot\vec{\sigma}) - a^2\vec{\sigma} \tag{2.5.46}$$

where $[A,B] = AB - BA$ is the commutator and $\{A,B\} = AB + BA$ is the anti-commutator.

Identities of order $(\vec{\sigma})^3$:

$$\vec{\sigma}\cdot((\vec{a}\cdot\vec{\sigma})\vec{\sigma}) = (\vec{\sigma}(\vec{a}\cdot\vec{\sigma}))\cdot\vec{\sigma} = -(\vec{a}\cdot\vec{\sigma}) \tag{2.5.47}$$

$$\vec{\sigma}\cdot(\vec{\sigma}(\vec{a}\cdot\vec{\sigma})) = ((\vec{a}\cdot\vec{\sigma})\vec{\sigma})\cdot\vec{\sigma} = 3(\vec{a}\cdot\vec{\sigma}) \tag{2.5.48}$$

$$(\vec{a}\cdot\vec{\sigma})\vec{\sigma}(\vec{a}\cdot\vec{\sigma}) = 2\vec{a}(\vec{a}\cdot\vec{\sigma}) - a^2\vec{\sigma} \tag{2.5.49}$$

Identity of order $(\vec{\sigma})^n$:

$$(\vec{a}\cdot\vec{\sigma})^n = \begin{cases} a^n & \text{n even} \\ a^{n-1}(\hat{a}\cdot\vec{\sigma}) & \text{n odd} \end{cases} \tag{2.5.50}$$

Finally, there are identities that relate to inner products taken with various forms of the spin vector. These identities are as follows:

$$\langle s\,|\vec{a}\cdot\vec{\sigma}|\,s\rangle = \vec{a}\cdot\langle s\,|\vec{\sigma}|\,s\rangle = \vec{a}\cdot\hat{s} \tag{2.5.51}$$

$$\langle s\,|\vec{a}\times\vec{\sigma}|\,s\rangle = \vec{a}\times\langle s\,|\vec{\sigma}|\,s\rangle = \vec{a}\times\hat{s} \tag{2.5.52}$$

$$\langle s\,|R\vec{\sigma}|\,s\rangle = R\langle s\,|\vec{\sigma}|\,s\rangle = R\hat{s} \tag{2.5.53}$$

where $\vec{a}$ and $\vec{b}$ are arbitrary vectors in Stokes space, $a$ is the length of $\vec{a}$, and $R$ is a 3x3 matrix. Identity (2.5.53) is always a source of confusion, so it is repeated explicitly:

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} \langle s\,|\sigma_1|\,s\rangle \\ \langle s\,|\sigma_2|\,s\rangle \\ \langle s\,|\sigma_3|\,s\rangle \end{pmatrix} = \begin{pmatrix} \langle s\,|r_{11}\sigma_1 + r_{12}\sigma_2 + r_{13}\sigma_3|\,s\rangle \\ \langle s\,|r_{21}\sigma_1 + r_{22}\sigma_2 + r_{23}\sigma_3|\,s\rangle \\ \langle s\,|r_{31}\sigma_1 + r_{32}\sigma_2 + r_{33}\sigma_3|\,s\rangle \end{pmatrix}$$

where $R\hat{s}$ on the left and $\langle s\,|R\vec{\sigma}|\,s\rangle$ on the right.

### 2.5.5 Conservation of Length

Expressions (2.5.27) and (2.5.30) complete the bilateral connection between Jones and Stokes vectors. Length must be conserved, of course, and this is verified now.

The Stokes-vector length is derived from the product $\hat{s}\cdot\hat{s} = s_1^2 + s_2^2 + s_3^2$. Consider one coordinate alone,

$$s_k^2 = \langle s\,|\sigma_k|\,s\rangle\langle s\,|\sigma_k|\,s\rangle \tag{2.5.54}$$

Substitution of the projector $|s\rangle\langle s|$, (2.5.29), for the innermost term gives

$$s_k^2 = \frac{1}{2}\,\langle s\,|s\rangle + \frac{1}{2}\,\langle s\,|\sigma_k(\hat{s}\cdot\vec{\sigma})\sigma_k|\,s\rangle \tag{2.5.55}$$

The sum of all three terms gives

$$s_1^2 + s_2^2 + s_3^2 = \frac{3}{2}\,\langle s\,|s\rangle + \frac{1}{2}\,\langle s\,|\vec{\sigma}\cdot((\hat{s}\cdot\vec{\sigma})\vec{\sigma})|\,s\rangle \tag{2.5.56}$$

The spin-vector identity (2.5.48) simplifies (2.5.56):

$$s_1^2 + s_2^2 + s_3^2 = \frac{3}{2}\,\langle s\,|s\rangle - \frac{1}{2}\,\langle s\,|\hat{s}\cdot\vec{\sigma}|\,s\rangle = \langle s\,|s\rangle \tag{2.5.57}$$

Thus, $\| \vec{s} \|^2 = \langle s | s \rangle$. Length is clearly preserved in this direction. For the reverse mapping, construction of (2.5.30) without the assumption $\langle s | s \rangle = 1$ produces

$$\vec{s} \cdot \vec{\sigma} \, | s \rangle = \langle s | s \rangle \, | s \rangle \quad \longrightarrow \quad \hat{s} \cdot \vec{\sigma} \, | s \rangle = | s \rangle \tag{2.5.58}$$

That length is conserved over the bilateral connections is thus established.

There is, however, one piece of information that is lost in the mapping from Jones to Stokes. Since Stokes coordinates are derived from intensity measurements, the common phase of the Jones vector is lost. Transformation from Stokes back to Jones does not reintroduce this phase. Any Jones vector constructed from a Stokes vector is accurate to the "true" Jones vector to within an arbitrary common phase. Physically this just means that the absolute time it took for the light to travel from its source to the observer cannot be determined from Stokes measurements.

### 2.5.6 Orthogonal Polarization States

For every polarization state $|s_+\rangle$ there is a unique polarization state $|s_-\rangle$ such that $\langle s_- | s_+ \rangle = 0$. These states $|s_+\rangle$ and $|s_-\rangle$ are orthogonal. Given $|s_+\rangle$ how does one can construct the orthogonal state $|s_-\rangle$ and its Stokes equivalent?

From (2.5.30) on page 56 one writes

$$\langle s_- | s_+ \rangle = \Big( \langle s_- | \, (\hat{s}_- \cdot \vec{\sigma})^\dagger \Big) \Big( (\hat{s}_+ \cdot \vec{\sigma}) \, | s_+ \rangle \Big) = 0 \tag{2.5.59}$$

Since $(\hat{s}_- \cdot \vec{\sigma})$ is Hermitian, (2.5.59) is rewritten using spin-vector identity (2.5.38) as

$$\langle s_- | s_+ \rangle = (\hat{s}_- \cdot \hat{s}_+) \langle s_- | s_+ \rangle + j \langle s_- | (\hat{s}_- \times \hat{s}_+) \cdot \vec{\sigma} | \, s_+ \rangle \tag{2.5.60}$$

As $\langle s_- | s_+ \rangle = 0$, (2.5.60) requires that $(\hat{s}_- \times \hat{s}_+) = 0$. There are two orientations that produce $(\hat{s}_- \times \hat{s}_+) = 0$: $\hat{s}_- \cdot \hat{s}_+ = \pm 1$. If $\hat{s}_- \cdot \hat{s}_+ = +1$, then $\langle s_- | s_+ \rangle = 1$, contradicting the orthogonality of the two states. Therefore it must be the case that

$$\hat{s}_- \cdot \hat{s}_+ = -1 \tag{2.5.61}$$

The Stokes coordinates for any two orthogonal polarization states are on opposite sides of the Poincaré sphere: $\hat{s}_- = -\hat{s}_+$. Specifically, a chord that connects any two orthogonal states crosses through the origin of the sphere. The polarimetric parameters $\chi$ and $\phi$ are related through the Stokes vectors as

$$\begin{pmatrix} \cos 2\chi_- \\ \sin 2\chi_- \cos \phi_- \\ \sin 2\chi_- \sin \phi_- \end{pmatrix} = - \begin{pmatrix} \cos 2\chi_+ \\ \sin 2\chi_+ \cos \phi_+ \\ \sin 2\chi_+ \sin \phi_+ \end{pmatrix} \tag{2.5.62}$$

A sufficient requirement to satisfy (2.5.62) is

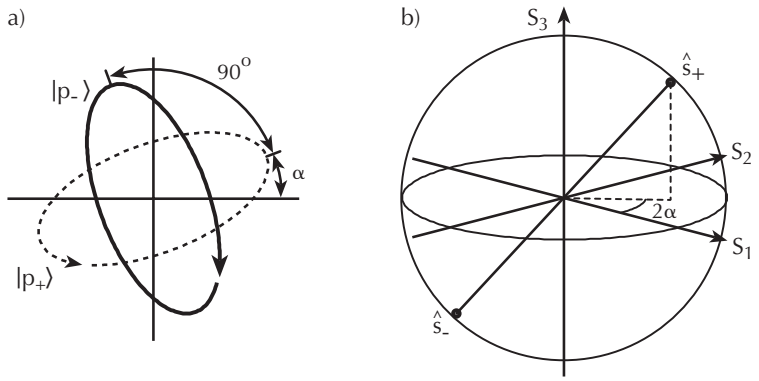


**Fig. 2.3.** Orthogonal polarization states in Jones and Stokes space. a) The handedness of the polarization ellipse is reversed and the major axis is rotated by $\pi/2$. b) Points on opposite sides of the Poincaré sphere are orthogonal.

$$2\chi_- = 2\chi_+ + \pi \tag{2.5.63a}$$

$$\phi_- = \phi_+ \tag{2.5.63b}$$

Equations (2.5.61) and (2.5.63) show that orthogonal polarization states have opposite handedness and perpendicular orientations of the respective elliptical major axes. The Jones and Stokes representation of orthogonal polarization pairs is illustrated in Fig. 2.3.

### 2.5.7 Non-Orthogonal Polarization States

The inner product magnitude between two polarization states may be calculated either in Jones or Stokes space. Consider two Jones vectors $|p\rangle$ and $|q\rangle$ that are not normalized, and recall that $\mathrm{Tr}\left(|p\rangle\langle q|\right) = \langle p|q\rangle$. In a manner similar to (2.5.29), the inner product between the two Jones vectors is written

$$|p\rangle\langle p| = \frac{1}{2}\left(I + \vec{p}\cdot\vec{\sigma}\right)\langle p|p\rangle \tag{2.5.64}$$

Multiplication on the right by $|q\rangle$ and on the left by $\langle q|$, and some rearrangement, makes

$$\frac{|\langle p|q\rangle|^2}{\langle p|p\rangle\langle q|q\rangle} = \frac{1}{2}\left(1 + \vec{p}\cdot\vec{q}\right) \tag{2.5.65}$$

When $|p\rangle$ and $|q\rangle$ are normalized, the identity reduces to

$$|\langle p|q\rangle|^2 = \frac{1}{2}\left(1 + \hat{p}\cdot\hat{q}\right) \tag{2.5.66}$$

The magnitude of the inner product in Jones space is derived directly from the Stokes vectors using this equation. What cannot be discerned, however, is

the phase of the inner product. To recover the phase, $\langle p\,|q\rangle$ must be calculated explicitly in Jones space. To construct the Jones vectors, one must either solve the eigenvalue equation (2.5.30) or make the Jones vector (1.4.19) on page 18 for both $|p\rangle$ and $|q\rangle$

### 2.5.8 Pauli Spin Operators

A general operator can be constructed from the identity matrix and spin-vector by the form

$$\begin{aligned} A &= a_0 I + \vec{a}\cdot\vec{\sigma} \\ &= a_0 I + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \\ &= \begin{pmatrix} a_0 + a_1 & a_2 - ja_3 \\ a_2 + ja_3 & a_0 - a_1 \end{pmatrix} \end{aligned} \tag{2.5.67}$$

where all $a_k$ are complex numbers. This matrix has the eight requisite independent variables necessary for a general Jones matrix. The entries in $A$ are isolated by the trace:

$$a_0 = \tfrac{1}{2}\,\mathrm{Tr}(A) \quad \text{and} \quad \vec{a} = \tfrac{1}{2}\,\mathrm{Tr}\,(A\vec{\sigma}) \tag{2.5.68}$$

A Hermitian operator is a special case of $A$:

$$H = a_0 I + \vec{a}\cdot\vec{\sigma}, \qquad a_k \text{ real} \tag{2.5.69}$$

The determinant is $\det(H) = a_0^2 - \left(a_1^2 + a_2^2 + a_3^2\right)$. Moreover, when the trace of $H$ is zero, the Hermitian matrix equals a spin-vector form

$$H_{\mathrm{Tr}=0} = \vec{a}\cdot\vec{\sigma} \tag{2.5.70}$$

Throughout much of this text, Hermitian operators with zero trace, and operators that preserve that trace, are associated with the spin-vector form.

The general operator $A$ can be decomposed into Hermitian and skew-Hermitian matrices

$$A = H_r + jH_i \tag{2.5.71}$$

where the operator $K = (jH_i)$ is skew-Hermitian: $K^\dagger = -K$. The eigenvalues of skew-Hermitian matrices are purely imaginary. The matrices $H_r$ and $H_i$ contain the real and imaginary parts of $A$, respectively. The decomposition is taken further by separating the finite-trace component from the traceless components. Writing the complex number $a_0$ as $a_0 = a_0' + ja_0''$ and identifying each traceless Hermitian matrix with a spin-vector form, one has

$$A = \vec{a}_r\cdot\vec{\sigma} + j\vec{a}_i\cdot\vec{\sigma} + (a_0' + ja_0'')\,I \tag{2.5.72}$$

This is the most general spin-vector form of an arbitrary operator $A$. In practice, the decomposition matrices are derived from $A$ as follows:

$$H_r = \tfrac{1}{2}\left(A + A^\dagger\right), \quad \text{and} \quad H_i = -\tfrac{j}{2}\left(A - A^\dagger\right) \tag{2.5.73}$$

The matrices $H_r$ and $H_i$ are made traceless by calculating $a_0 = \frac{1}{2}\mathrm{Tr}(H)$ for the real and imaginary components. The real-valued Stokes vectors $\vec{a}_r$ and $\sigma_i$ can be read from the matrices $\vec{a}_{r,i} \cdot \vec{\sigma} = H_{r,i} - a_{r,i,0} I$.

The Pauli spin operator $\mathcal{S}$ produces the most compact way to describe operators and concatenations of operators. The spin operator is a matrix exponential form of $A$ and can describe Hermitian, unitary, and general operators. The matrix exponential is written

$$\mathcal{S} = \exp\left(A/2\right) \tag{2.5.74}$$

The exponential is evaluated using its Taylor expansion. For instance,

$$\begin{aligned}\exp(M) &= I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \frac{1}{4!}M^4 + \ldots \\ &= \left(I + \frac{1}{2!}M^2 + \frac{1}{4!}M^4 + \ldots\right) + \left(M + \frac{1}{3!}M^3 + \frac{1}{5!}M^5 + \ldots\right)\end{aligned} \tag{2.5.75}$$

For a Hermitian matrix, $M = \alpha_o I + (\vec{\alpha} \cdot \vec{\sigma})$ where $\alpha_k$ are real. The $n^{\text{th}}$-order spin-vector identity (2.5.50) gives the necessary reductions for $(\vec{\alpha} \cdot \vec{\sigma})^n$, which gives

$$\exp\left(\vec{\alpha} \cdot \vec{\sigma}/2\right) = I \cosh\left(\alpha/2\right) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh\left(\alpha/2\right) \tag{2.5.76}$$

where $\vec{\alpha} = \alpha\hat{\alpha}$. The Pauli spin operator for a Hermitian operator is thus

$$\begin{aligned}\mathcal{S}_H &= \exp\left(\alpha_0/2\right) \exp\left(\vec{\alpha} \cdot \vec{\sigma}/2\right) \\ &= \mathrm{e}^{\alpha_0/2} \left[\, I \cosh\left(\alpha/2\right) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh\left(\alpha/2\right) \,\right]\end{aligned} \tag{2.5.77}$$

One can interpret this operator as that for a partial polarizer: the common loss is $\alpha_0/2$ (which is a negative quantity for loss), the maximum and minimum differential losses are $1 \pm \tanh \alpha/2$, and the Stokes direction of partial polarization is $\hat{\alpha}$.

The Pauli spin operator for a unitary matrix is constructed by recalling the connection between Hermitian and unitary operators (2.4.18) on page 49. For coefficients $\beta_k$ real, the unitary form of $M$ is $M = -j(\beta_o I + (\vec{\beta} \cdot \vec{\sigma}))$. The equivalent to (2.5.76) is

$$\exp\left(-j\,\vec{\beta} \cdot \vec{\sigma}/2\right) = I \cos\left(\beta/2\right) - j(\hat{\beta} \cdot \vec{\sigma}) \sin\left(\alpha/2\right) \tag{2.5.78}$$

where $\vec{\beta} = \beta\hat{\beta}$. The Pauli spin operator for a unitary operator is thus

$$\begin{aligned}\mathcal{S}_T &= \exp\left(-j\,\beta_0/2\right)\exp\left(-j\vec{\beta}\cdot\vec{\sigma}/2\right)\\ &= \mathrm{e}^{-j\,\beta_0/2}\,\left[\,I\cos\left(\beta/2\right)-j(\hat{\beta}\cdot\vec{\sigma})\sin\left(\beta/2\right)\,\right]\end{aligned}\tag{2.5.79}$$

One interprets this operator as a retardation plate: $-\beta_0/2$ is the common phase, the full retardance is $\beta$, and the axis of retardation is $\hat{\beta}$. When the common phase is removed the unitary matrix $U$ is recovered.

In the particular case where the axes of polarization and retardance coincide, the compound effect is expressed as the spin vector $(-j\vec{\beta}+\vec{\alpha})\cdot\vec{\sigma}$. Following the form of (2.5.76) the Pauli spin operator expands to matrix form as

$$\begin{aligned}&\exp\left((-j\,\beta_0+\alpha_0)/2\right)\exp\left((-j\vec{\beta}+\vec{\alpha})\cdot\vec{\sigma}/2\right)=\\ &\mathrm{e}^{(-j\beta_0+\alpha_0)/2}\,\left[\,I\cosh\left((-j\beta+\alpha)/2\right)+(\hat{r}\cdot\vec{\sigma})\sinh\left((-j\beta+\alpha)/2\right)\,\right]\end{aligned}\tag{2.5.80}$$

where $\hat{r}$ is the axis of polarization and retardance.

Two relevant properties of matrix exponentials are

$$\frac{\partial}{\partial t}\,\mathrm{e}^{At}=A\mathrm{e}^{At}=\mathrm{e}^{At}A\tag{2.5.81}$$

$$\frac{\partial}{\partial t}\,\mathrm{e}^{At}\mathrm{e}^{Bt}\mathrm{e}^{Ct}=A\mathrm{e}^{At}\mathrm{e}^{Bt}\mathrm{e}^{Ct}+\mathrm{e}^{At}B\mathrm{e}^{Bt}\mathrm{e}^{Ct}+\mathrm{e}^{At}\mathrm{e}^{Bt}\mathrm{e}^{Ct}C\tag{2.5.82}$$

while two non-intuitive results are

$$\frac{\partial}{\partial A}\,\mathrm{e}^{At}\neq t\mathrm{e}^{At}\tag{2.5.83}$$

$$\mathrm{e}^{At}\mathrm{e}^{Bt}\neq\mathrm{e}^{(A+B)t}\tag{2.5.84}$$

unless, for the second equation, $A$ and $B$ commute.

When one writes a compound operation such as

$$H_1UH_2=\exp\left((\vec{\alpha}_1\cdot\vec{\sigma})/2\right)\exp(-j(\vec{\beta}\cdot\vec{\sigma})/2)\exp\left((\vec{\alpha}_2\cdot\vec{\sigma})/2\right)\tag{2.5.85}$$

one is only figuratively expressing the series of operations on a polarization state. The evaluation of $H_1UH_2$ requires substitution of the matrix form for each operator.

## 2.6 Equivalent Unitary Transformations

The preceding sections have established the bilateral connection between Jones and Stokes vectors, and have constructed Hermitian, unitary, and general Jones operators that act on Jones vectors. The connection between a Hermitian matrix and an equivalent Stokes matrix is made with the Mueller

matrix, (1.4.22) on page 18 and figuratively (2.1.3) on page 38. Mueller matrices operate on $4 \times 1$ Stokes vectors to create new, transformed $4 \times 1$ Stokes vectors.

The connection between a unitary matrix and an equivalent Stokes matrix is also made with the Mueller matrix, but as indicated by (2.1.1) on page 38, only the lower right $3 \times 3$ sub-matrix $\mathbf{R}$ is relevant. Sub-matrix $\mathbf{R}$ maps spherical coordinates $(S_1, S_2, S_3)$ into new spherical coordinates $(S_1', S_2', S_3')$ without change of length. Therefore one expects the existence of a rotation operator $R$ corresponding to matrix $\mathbf{R}$ that performs rotations on the Poincaré sphere. The operator $R$ does indeed exist and its derivation and properties are so central to the description of retardance that this entire section is devoted to its understanding.

Operators $U$ and $R$ are equivalent representation of the same transformation cast in two different vector spaces. The operators are called isomorphic because they have similar, but not equal, effects. The isomorphism is two-to-one since, as well be seen, there are two Jones operations that have the same effect as every one Stokes operation.

Consider equivalent vectors $|s\rangle$ and $\hat{s}$ at the input of a system and equivalent vectors $|t\rangle$ and $\hat{t}$ at the output. In Jones space a unitary transformation $T$, corresponding to the underlying Maxwell's equations in anisotropic media, links the input and output. In Stokes space the rotation operator $R$ links the input and output. The parallel transformations are

$$|t\rangle = T\,|s\rangle \quad \xleftrightarrow{\text{dual}} \quad \hat{t} = R\,\hat{s} \tag{2.6.1}$$

Expansion of the Stokes vectors on the right side of (2.6.1) into their corresponding inner products gives the relation between $R$ and $T$

$$\langle t\,|\vec{\sigma}|\,t\rangle = R\langle s\,|\vec{\sigma}|\,s\rangle \tag{2.6.2}$$

Replacing $|t\rangle$ with $T\,|s\rangle$ and applying identity (2.5.53) gives

$$\langle s\,\big|T^\dagger \vec{\sigma} T\big|\,s\rangle = \langle s\,|R\,\vec{\sigma}|\,s\rangle \tag{2.6.3}$$

Since (2.6.3) holds for any $|s\rangle$, the embedded operators must be equal. Therefore

$$R\,\vec{\sigma} = U^\dagger \vec{\sigma} U \tag{2.6.4}$$

where the common phase of $T$ commutes with $\vec{\sigma}$ and is eliminated. Equation (2.6.4) has an unusual form; the interpretation is

$$\begin{pmatrix} \\ \quad 3\times 3 \quad \\ \\ \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} U^\dagger \sigma_1 U \\ U^\dagger \sigma_2 U \\ U^\dagger \sigma_3 U \end{pmatrix} \tag{2.6.5}$$

That $R$ is unitary is derived by multiplying (2.6.4) by its adjoint:

$$\vec{\sigma}^\dagger R^\dagger R\,\vec{\sigma} = U^\dagger \vec{\sigma}^\dagger U U^\dagger \vec{\sigma} U \tag{2.6.6}$$

The product $\vec{\sigma}^\dagger \vec{\sigma} = 3I$, so only if $R^\dagger R = I$ does (2.6.6) hold. Therefore,

$$R^\dagger R = R R^\dagger = I \tag{2.6.7}$$

Since $R$ is unitary it embodies a pure rotation; there is no scaling or translation; $\hat{t}$ is related to $\hat{s}$ by a rotation in Stokes space. The group properties of $R$ and their correspondence to the group properties of $U$ are listed in the following section.

There are two ways to derive an explicit expression for $R$, either through the matrix form of $U$, generating a matrix form of $R$; or through the Pauli spin operator form of $U$, generating a vector form of $R$. The vector form is more "lightweight" and powerful than the matrix form because successive operations in Stokes space are evaluated purely in vector form without an *a priori* choice of orthonormal basis. The vector form is also a template with which to construct any matrix $R$ without matrix multiplication.

### 2.6.1 Group Properties of SU(2) and O(3)

The group defined by multiplication operators of $R$ is called O(3), where O stands for orthogonal and (3) for the three rotational dimensions of $R$. The group properties for multiplication mirror those for SU(2), cf. §2.4.6:

| | | |
|---|---|---|
| Identity | $UI = U$ | $RI = R$ |
| Closure | $U_1 U_2 = U_3$ | $R_1 R_2 = R_3$ |
| Inverse | $U^{-1}U = I$ | $R^{-1}R = I$ |
| Associativity | $(U_1U_2)U_3 = U_1(U_2U_3)$ | $(R_1R_2)R_3 = R_1(R_2R_3)$ |

where in all cases $U_{1,2,3} \in \mathrm{SU}(2)$ and $R_{1,2,3} \in \mathrm{O}(3)$.

To confirm the entries in the above table, note that multiplication of successive $U$ and $R$ operators form a one-to-one correspondence. For example,

$$|t\rangle = T_2 T_1 |s\rangle \quad \overset{\text{dual}}{\longleftrightarrow} \quad \hat{t} = R_2 R_1 \hat{s} \tag{2.6.8}$$

Operator equality is generated through the expression

$$R_2 R_1 \vec{\sigma} = {U_1}^\dagger {U_2}^\dagger \vec{\sigma} U_2 U_1 \tag{2.6.9}$$

Making the correspondences $R' = R_2 R_1$ and $U' = U_2 U_1$, then (2.6.9) is equivalent to (2.6.4). Given that $U_2 U_1 \in \mathrm{SU}(2)$ and $U' \longleftrightarrow R'$, one infers closure on O(3).

There are twice as many entries in the group SU(2) as in O(3). For every $U_o$, the corresponding $R_o$ is $R_o \vec{\sigma} = {U_o}^\dagger \vec{\sigma} U_o$. For every $-U_o$ the corresponding operator $R_o$ is the same: $R_o \vec{\sigma} = (-U_o)^\dagger \vec{\sigma} (-U_o)$. The isomorphism between SU(2) and O(3) is two-to-one.

### 2.6.2 Matrix Entries of $R$ in a Fixed Coordinate System

One way to generate $R$ explicitly is through the matrix form of $U$. Earlier, the generation of matrix entries for the Mueller matrix from entries in the Jones matrix was given without derivation, (1.4.22) on page 18. That equation created a $4 \times 4$ matrix; only for a unitary matrix is a distinct $3 \times 3$ sub-matrix formed. The matrix entries for $\mathbf{M}$ are found as follows.

Start with the two relations $\hat{t}_j = \langle t \,|\sigma_j|\, t \rangle$ and $\hat{t}_j = \mathrm{Tr}\,(M\sigma_j)$, equations (2.5.24) on page 55 and (2.5.68) on page 61, respectively. The index $j$ is for $j = 0, 1, 2, 3$. Moreover, assume the system $|t\rangle = J\,|s\rangle$. Identification with $\langle t\,|t\rangle = \mathrm{Tr}(|t\rangle\langle t|)$ gives

$$\begin{aligned} t_j &= \langle t \,|\sigma_j|\, t \rangle \\ &= \mathrm{Tr}\,(|t\rangle\langle t|\,\sigma_j) \\ &= \mathrm{Tr}\left(J\,|s\rangle\langle s|\,J^\dagger\sigma_j\right) \end{aligned} \tag{2.6.10}$$

Next, the outer product $|s\rangle\langle s|$ is replaced with something close to the spin-vector form (2.5.29) on page 56. Since the vector is not necessarily normalized, the expression $|s\rangle\langle s| = \frac{1}{2}\langle s\,|s\rangle\,(I + \hat{s}\cdot\vec{\sigma})$ will be used. Thus,

$$\begin{aligned} \mathrm{Tr}\left(J\,|s\rangle\langle s|\,J^\dagger\sigma_j\right) &= \tfrac{1}{2}\,\langle s\,|s\rangle\;\mathrm{Tr}\left(J\,(I + \hat{s}\cdot\vec{\sigma})\,J^\dagger\sigma_j\right) \\ &= \tfrac{1}{2}\,\langle s\,|s\rangle\;\mathrm{Tr}\left(J(\hat{s}\cdot\vec{\sigma})J^\dagger\sigma_j\right) \\ &= \frac{1}{2}\sum_{k=0}^{3}\mathrm{Tr}\left(J\sigma_k J^\dagger\sigma_j\right)\;s_k \end{aligned} \tag{2.6.11}$$

where $\vec{\sigma}$ has been loosely indexed here to include $\sigma_0$, $s_k = \langle s\,|s\rangle\,\hat{s}_k$, and the common phase in $T$ commutes with $(\hat{s}\cdot\vec{\sigma})$ and is eliminated. The last equation has the matrix multiplication form

$$t_j = \sum_{k=0}^{3} M_{j+1,k+1}s_k \tag{2.6.12}$$

Identification of the matrix entries gives the final expression

$$M_{j+1,k+1} = \frac{1}{2}\,\mathrm{Tr}\left(J\sigma_k J^\dagger\sigma_j\right) \tag{2.6.13}$$

The specialized case of $J = U$ is of immediate interest. Substitution of $U$ for $J$ in (2.6.13) shows that for $k = 0, j \neq 0$, and for $j = 0, k \neq 0$, the matrix entries are identically zero. This is because $\mathrm{Tr}(\sigma_j) = 0$. Other than the $M_{1,1}$ matrix entry, which is unity, only the sub-matrix $R$ survives the trace. Explicitly, the matrix entries for $R$ given a matrix form of $U$ are

$$R_{j,k} = \frac{1}{2}\,\mathrm{Tr}\left(U\sigma_k U^\dagger\sigma_j\right) \tag{2.6.14}$$

**Table 2.1.** Elementary Rotations in Jones and Stokes Space

| | | |
|---|---|---|
| $\kappa = 0$ | $U = \begin{pmatrix} \mathrm{e}^{j\alpha} & \\ & \mathrm{e}^{-j\alpha} \end{pmatrix}$ | $R_1 = \begin{pmatrix} 1 & & \\ & \cos 2\alpha & \sin 2\alpha \\ & -\sin 2\alpha & \cos 2\alpha \end{pmatrix}$ |
| $\alpha = 0$<br>$\beta = \pi/2$ | $U = \begin{pmatrix} \cos\kappa & -j\sin\kappa \\ -j\sin\kappa & \cos\kappa \end{pmatrix}$ | $R_2 = \begin{pmatrix} \cos 2\kappa & & \sin 2\kappa \\ & 1 & \\ -\sin 2\kappa & & \cos 2\kappa \end{pmatrix}$ |
| $\alpha = \beta = 0$ | $U = \begin{pmatrix} \cos\kappa & -\sin\kappa \\ \sin\kappa & \cos\kappa \end{pmatrix}$ | $R_3 = \begin{pmatrix} \cos 2\kappa & -\sin 2\kappa & \\ \sin 2\kappa & \cos 2\kappa & \\ & & 1 \end{pmatrix}$ |

Using the general form of $U$ given in (2.4.26) on page 51, $R$ is resolved as

$$R = \begin{pmatrix} \cos 2\kappa & -\cos(\alpha-\beta)\sin 2\kappa & -\sin(\alpha-\beta)\sin 2\kappa \\ \cos(\alpha+\beta)\sin 2\kappa & \cos 2\alpha\cos^2\kappa - \cos 2\beta\sin^2\kappa & \sin 2\alpha\cos^2\kappa + \sin 2\beta\sin^2\kappa \\ -\sin(\alpha+\beta)\sin 2\kappa & -\sin 2\alpha\cos^2\kappa + \sin 2\beta\sin^2\kappa & \cos 2\alpha\cos^2\kappa + \cos 2\beta\sin^2\kappa \end{pmatrix} \tag{2.6.15}$$

Calculation of the determinant gives

$$\det(R) = 1 \tag{2.6.16}$$

which verifies that $R$ is invertible, unitary, and contains no reflections.

Table 2.1 gives the three elementary rotations about the Stokes axes $(S_1, S_2, S_3)$ and their original unitary form. Notice that all angles which appear in the unitary matrices are doubled in the corresponding Stokes matrices. Stokes angles are twice that of physical, or "laboratory" angles. This is also why the Jones to Stokes isomorphism is two-to-one: rotation of $\pi$ in Jones space is invariant, and so is rotation by $2\pi$ in Stokes space.

### 2.6.3 Vector Expression of $R$ in a Local Coordinate System

The vector form of $R$ abstracts away any notion of an underlying, fixed coordinate system. Rather, each operation $R$ has its own local coordinate system based on the eigenvectors and spin direction of $R$. The vectorial form of $R$ gives the highest level of geometric interpretation to transformation mechanics in Stokes space.

The vector expression for $R$ is derived from the vector form of $U$. The operator $U$ is resolved into its eigenvector-based projectors using (2.4.15) on page 48; the resolution for a two-dimensional system gives

$$U = \mathrm{e}^{-j\varphi/2}\,|r_+\rangle\langle r_+| + \mathrm{e}^{j\varphi/2}\,|r_-\rangle\langle r_-| \tag{2.6.17}$$

where the projectors are equated to the spin-vector form via

$$|r_\pm\rangle\langle r_\pm| = \frac{1}{2}\,(I \pm \hat{r}\cdot\vec{\sigma}) \tag{2.6.18}$$

Note that $\hat{r}_+ = -\hat{r}_-$ are orthogonal Stokes coordinates. In the following, $\hat{r}$ will denote the eigenvector with the positive subscript. Substitution of the spin-vector form of the projector into $U$ gives

$$U = I\cos(\varphi/2) - j(\hat{r}\cdot\vec{\sigma})\sin(\varphi/2) \tag{2.6.19}$$

Equation (2.6.19) is now in familiar form and can be mapped to the exponential equivalent as

$$U = \mathrm{e}^{-j(\varphi/2)(\hat{r}\cdot\vec{\sigma})} \tag{2.6.20}$$

where $(\varphi/2)(\hat{r}\cdot\vec{\sigma})$ is the Hermitian operator associated with $U$.

Substitution of (2.6.19) into the equivalence relation (2.6.4), and applying the spin-vector identities (2.5.35), (2.5.36), and (2.5.49) produces

$$R\,\vec{\sigma} = \cos\varphi\vec{\sigma} + (1-\cos\varphi)\hat{r}(\hat{r}\cdot\vec{\sigma}) + \sin\varphi(\hat{r}\times\vec{\sigma}) \tag{2.6.21}$$

Since each term on the left- and right-hand sides of (2.6.21) operates on $\vec{\sigma}$, one can extract the embedded relation for $R$

$$R = \cos\varphi I + (1-\cos\varphi)(\hat{r}\hat{r}\cdot) + \sin\varphi(\hat{r}\times) \tag{2.6.22}$$

Grouping like terms gives

$$R = (\hat{r}\hat{r}\cdot) + \sin\varphi(\hat{r}\times) + \cos\varphi(I - (\hat{r}\hat{r}\cdot)) \tag{2.6.23}$$

Recalling the vector identity $\mathbf{a}\times\mathbf{a}\times\mathbf{c} = \mathbf{b}(\mathbf{a}\cdot\mathbf{a}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})$, the last term on the right-hand side is identified as

$$(\hat{r}\times)(\hat{r}\times) = (\hat{r}\hat{r}\cdot) - I \tag{2.6.24}$$

The final vectorial form of $R$ is then

$$R = (\hat{r}\hat{r}\cdot) + \sin\varphi(\hat{r}\times) - \cos\varphi(\hat{r}\times)(\hat{r}\times) \tag{2.6.25}$$

Equation (2.6.25) is a beautifully compact expression for the action any unitary operator has on a polarization vector. The vector $\hat{r}$ points in the direction of the positive eigenvector of $U$. The vector operators $\{(\hat{r}\hat{r}\cdot), (\hat{r}\times), (\hat{r}\times)(\hat{r}\times)\}$ form a local orthonormal basis. The local basis requires a vector about which

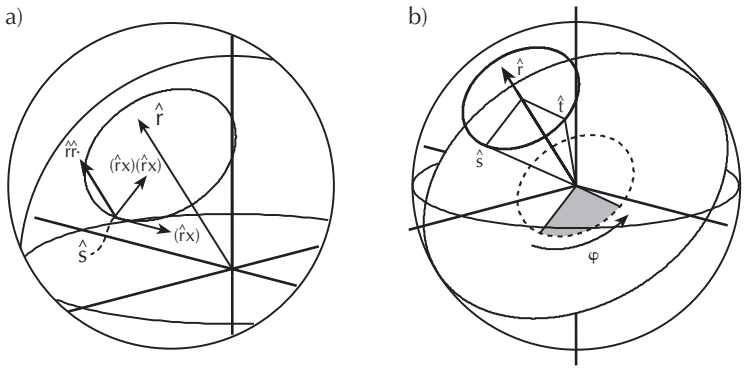


**Fig. 2.4.** Vector components of rotation operator $R$. a) The local orthonormal basis $\{(\hat{r}\hat{r}\cdot), (\hat{r}\times), (\hat{r}\times)(\hat{r}\times)\}$ as resolved on $\hat{s}$. b) Transformation to $\hat{t}$ from $\hat{s}$ via precession about $\hat{r}$, travelling through precession angle $\varphi$.

the basis can be fully resolved; for instance, operation on state $\hat{s}$ generates the basis $(\hat{r}, \hat{r} \times \hat{s}, \hat{r} \times \hat{r} \times \hat{s})$. In the absence of being fully resolved, the local basis has immutable properties that are independent of the resolving vector.

Figure 2.4(a) illustrates the local basis resolved by $\hat{s}$. Vector $\hat{s}$ in relation to $\hat{r}$ defines a precession circle, the circle about which $\hat{s}$ travels. Local axis $(\hat{r}\hat{r}\cdot)$ always points parallel to $\hat{r}$. The local axes $(\hat{r}\times), (\hat{r}\times)(\hat{r}\times)$ define the plane of the precession circle and are perpendicular to $\hat{r}$. The local axis $(\hat{r}\times)$ is tangent to the precession circle and $(\hat{r}\times)(\hat{r}\times)$ points to the origin of the precession circle. The particular pointing directions of $(\hat{r}\times)$ and $(\hat{r}\times)(\hat{r}\times)$, while always in the precession plane, are determined only after determination of $\hat{s}$. Figure 2.4(b) illustrates transformation to state $\hat{t}$ from $\hat{s}$ about $\hat{r}$. The precession angle is $\varphi$ and the precession direction follows the right-hand rule.

Since the motion of precession is so central in the description of polarization transformation mechanics, Fig. 2.5 is included to describe precession in a local coordinate system. Consider the input state $\hat{s}$ and the precession axis $\hat{r}$. The precession axis can be the birefringent axis of a dielectric medium or the principal-state-of-polarization axis used to describe polarization-mode dispersion. In any case, the angle $\gamma$ separates the two vectors. The motion of precession is to turn $\hat{s}$ about $\hat{r}$ in a circle while keeping the angle $\gamma$ fixed. This is the same motion a gyroscope exhibits under gravitational influence. The angle subtended by projections of states $\hat{s}$ and $\hat{t}$ onto the base circle is the precession angle. The differential equation of motion can be deduced from $R$ in local-coordinate form. Consider state $\hat{s}$ that undergoes a small change in angle $\delta\varphi$. The motion is

$$\hat{s} + \delta\hat{s} = R_{\delta\varphi}\hat{s} \tag{2.6.26}$$

Taking $R$ in the form (2.6.22) and simplifying for small angles,

$$\hat{s} + \delta\hat{s} = I\hat{s} + \delta\varphi\, \hat{r} \times \hat{s} \tag{2.6.27}$$

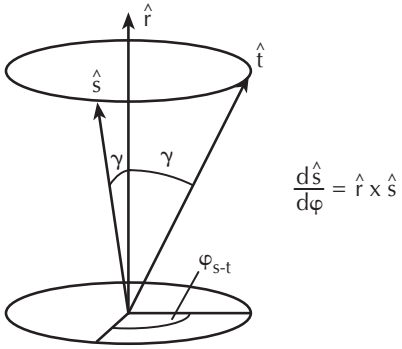


**Fig. 2.5.** Precessional motion of $\hat{s}$ about $\hat{r}$, passing through state $\hat{t}$. Angle $\gamma$ remains fixed, while angle $\varphi$, as projected onto the base, is the degree of precession.

which is rewritten in differential form as

$$\frac{d\hat{s}}{d\varphi} = \hat{r} \times \hat{s} \tag{2.6.28}$$

The $\hat{r} \times \hat{s}$ term dictates that $\hat{s}$ moves perpendicular to $\hat{r}$.

Equation 2.6.22 can be used as a template to construct a matrix representation for $R$ given any $\hat{r}$ and $\varphi$. The matrix representations for $\hat{r}\hat{r}\cdot$ and $\hat{r}\times$ are

$$(\hat{r}\hat{r}\cdot) = \begin{pmatrix} r_1 r_1 & r_1 r_2 & r_1 r_3 \\ r_2 r_1 & r_2 r_2 & r_2 r_3 \\ r_3 r_1 & r_3 r_2 & r_3 r_3 \end{pmatrix}, \quad (\hat{r}\times) = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix} \tag{2.6.29}$$

The first matrix is a projector, as verified by $\det(\hat{r}\hat{r}\cdot) = 0$. The second matrix is derived from $\hat{r}\times = \hat{r} \times S_1 + \hat{r} \times S_2 + \hat{r} \times S_3$, where $S_{1,2,3}$ are the three Stokes axes.

### 2.6.4 Select Vector Identities

Figure 2.6 illustrates two identities that have a simple geometric interpretation. The identity

$$\hat{r} \cdot (R\,\vec{a}\,) = \hat{r} \cdot \vec{a} \tag{2.6.30}$$

is illustrated by Fig. 2.6(a). The rotation of $\vec{a}$ generated by $R$ about $\hat{r}$ through angle $\varphi$ does not change the angle between $\hat{r}$ and $\vec{a}$. Thus the dot product may be taken with or without the intermediate rotation. The identity

$$(R\,\vec{a}\,) \cdot (R\,\vec{b}\,) = \vec{a} \cdot \vec{b} \tag{2.6.31}$$

is illustrated by Fig. 2.6(b). The rotation due to $R$ does not change the angle between vectors $\vec{a}$ and $\vec{b}$, so again the dot product may be taken with or without the intermediate rotation.

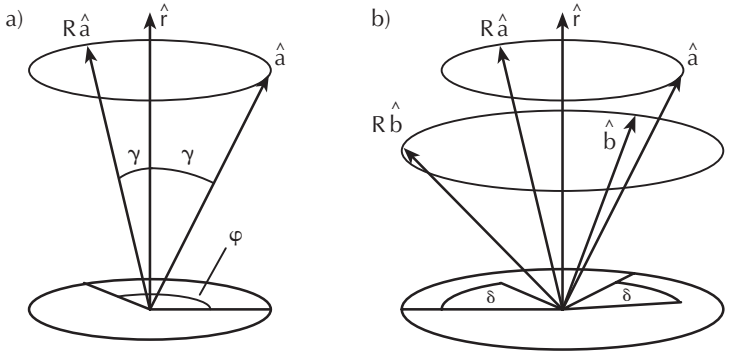


**Fig. 2.6.** Geometric representation of two rotational identities. a) $\hat{r} \cdot (R\,\vec{a}\,) = \hat{r} \cdot \vec{a}$. b) $(R\,\vec{a}\,) \cdot (R\,\vec{b}\,) = \vec{a} \cdot \vec{b}$.

### 2.6.5 Euler Rotations

The Euler rotations are an alternative method to construct the operator $R$ in matrix form (2.6.15). While there are several ways to establish the connection, the fact that multiplication in Stokes space corresponds to multiplication in Jones space is simple enough to construct the operator $U$ in the form (2.4.26), and to map the terms from Jones back to Stokes space. One can verify that

$$U = \begin{pmatrix} \mathrm{e}^{ju} & \\ & \mathrm{e}^{-ju} \end{pmatrix} \begin{pmatrix} \cos\kappa & -\sin\kappa \\ \sin\kappa & \cos\kappa \end{pmatrix} \begin{pmatrix} \mathrm{e}^{jv} & \\ & \mathrm{e}^{-jv} \end{pmatrix} \tag{2.6.32}$$

where $u = (\alpha + \beta)/2$ and $v = (\alpha - \beta)/2$. Identification of the Jones operators in (2.6.32) with the Stokes operators in Table 2.1 gives the equivalent Stokes transformations

$$R = R_1(\alpha + \beta)\; R_3(2\kappa)\; R_1(\alpha - \beta) \tag{2.6.33}$$

These rotations generate the general rotation matrix in (2.6.15).

Another way to view the unitary operator is to diagonalize the matrix. The eigenvalues of $U$ are complex exponentials that have unity magnitude and they are conjugates of one another: $U\,|r_\pm\rangle = \exp(\mp j\varphi/2)\,|r_\pm\rangle$. The operator $U$ can be separated as

$$U = V \Lambda V^\dagger \tag{2.6.34}$$

where the matrix $V$ is a $2 \times 2$ unitary matrix with the two eigenvectors $|r_\pm\rangle$ of $U$ entered as columns of $V$, and where the matrix $\Lambda$ is a diagonal $2 \times 2$ matrix with the eigenvalues of $U$ on the diagonal. The corresponding Stokes operation is

$$R = R_{\mathrm{E}} R_1(\varphi) {R^\dagger}_{\mathrm{E}} \tag{2.6.35}$$

where $R_{\mathrm{E}}$ is the Euler rotation associated with $V$. Now, several observations can be made. If the state at the input is $|r_\pm\rangle$, then the action of $U$ does not

change the state, only a phase is contributed. The state is invariant under $U$. For every $|r_\pm\rangle$ there are corresponding $\hat{r}_\pm$ vectors in Stokes space. The behavior of $R^\dagger{}_{\rm E}$ is to rotate $\hat{r}_\pm$ to $\pm s_1$; $R_1(\varphi)$ then pirouettes the state about the $s_1$ axis, and $R_{\rm E}$ returns the state back to $\hat{r}_\pm$.

The decomposition of $U$ as in (2.6.34) has much significance in relation to propagation through birefringent media. For example, the propagation constants for ordinary and extraordinary waves in a birefringent medium are $\beta_o = \omega n_o/c$ and $\beta_e = \omega n_e/c$. The eigenvalues of the propagation matrix are $\exp(\mp j(\beta_e - \beta_o)z/2)$. The polarization transformation in Stokes space is accordingly

$$R_{\Delta\beta} = R_{\rm E} R_x(\Delta\beta z) R^\dagger{}_{\rm E} \tag{2.6.36}$$

The inner matrix $R_x$ creates precession about the $s_1$ axis in Stokes space while the Euler rotation and its adjoint transforms the eigenstates of the system onto the $s_1$ axis and then restores the pointing direction of the eigenstate. The precession about $s_1$ is transformed to precession about $\hat{r}$.

### 2.6.6 Some Relevant Transformation Applications

Four examples are presented here to give some illustrative detail on how to use the Stokes transformation operator $R$ cast in local-coordinate form. First, differential precession rules for a single homogeneous birefringent section are written. Then, the polarization evolution through a concatenation of two misaligned birefringent sections, as a function of length, is illustrated. Third, the shortest distance between two polarization states is found. Finally, uniform and biased polarization scattering examples are given.

For polarization studies, the axis $\hat{r}$ is the extraordinary axis of a birefringent medium. For a birefringent crystal the birefringent axis lies in the equatorial plane in Stokes space. An input state $\hat{s}$ precesses about $\hat{r}$ as the state propagates through the medium. The retardation of a birefringent plate is just the angle $\varphi$ through which the state processes: $\varphi = \omega\Delta n L/c$, albeit any two rotations of $\varphi$ that are the same modulo $2\pi$ yield identical Stokes transformations. The angle $\varphi$ is called the birefringent phase.

The differential equation of motion for a state of polarization as it evolves through a homogeneous birefringent material is the same for differentials in either position or frequency. The retardation as a function of length $z$ and radial frequency $\omega$ is

$$\varphi = \frac{\omega\Delta n z}{c} \tag{2.6.37}$$

Moreover, the birefringent axis $\Omega$ is parallel to $\hat{r}$. Differentiation of (2.6.37) with respect to $z$ and substitution into (2.6.28) gives

$$\frac{d\hat{s}}{dz} = \Omega \times \hat{s} \tag{2.6.38}$$

where $\Omega = (\omega\Delta n/c)\hat{r}$. Equally possible is differentiation with respect to $\omega$, which gives

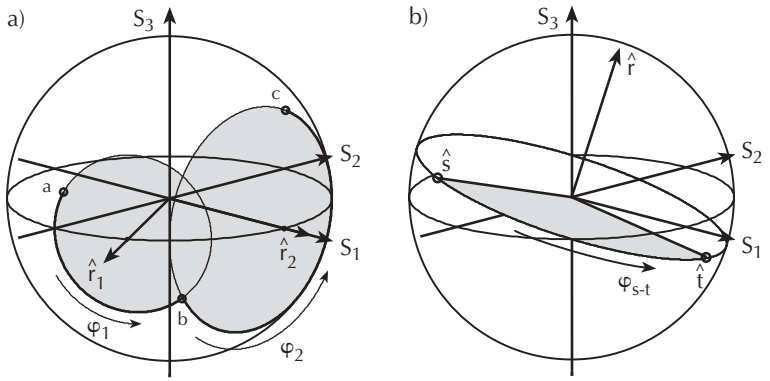


**Fig. 2.7.** a) Polarization evolution through two misaligned birefringent sections. Input state (a) precesses about $\hat{r}_1$ to state (b). That state enters the second stage, precesses about $\hat{r}_2$, and leaves as state (c). b) Construction of great circle through states $\hat{s}$ and $\hat{t}$. The normal to the circle is $\hat{r}$.

$$\frac{d\hat{s}}{d\omega} = \Omega \times \hat{s} \tag{2.6.39}$$

The differential precession rule for a single homogeneous birefringent section is the same whether the position or frequency changes. This simplicity is quickly broken when two or more homogeneous sections are concatenated. Birefringent concatenation is in the category of polarization-mode dispersion.

The polarization state evolution through two misaligned birefringent sections as a function of length can be evaluated using (2.6.25) in the following way. Since the media are misaligned, their birefringent axes are not parallel; that is, $\hat{r}_1 \neq \hat{r}_2$. Figure 2.7(a) illustrates the polarization evolution through the sections. The input state, arbitrarily selected, is located at position (a). That state precesses about $\hat{r}_1$ through angle $\varphi_1$, dictated by the length and birefringence of the section, as well as the input frequency. The output polarization from the first section is located at position (b). That state then enters the second section which transforms it about $\hat{r}_2$. The polarization state now traces a second circle that is different from the first. The output state is eventually located at position (c). The aggregate polarization transformation is calculated by the concatenation of $R_2$ and $R_1$. The compounded polarization transformation is

$$R_2 R_1 = \Big( (\hat{r}_2\hat{r}_2\cdot) + \sin\varphi_2(\hat{r}_2\times) - \cos\varphi_2(\hat{r}_2\times)(\hat{r}_2\times) \Big) \tag{2.6.40}$$

$$\Big( (\hat{r}_1\hat{r}_1\cdot) + \sin\varphi_1(\hat{r}_1\times) - \cos\varphi_1(\hat{r}_1\times)(\hat{r}_1\times) \Big) \tag{2.6.41}$$

Each transformation is denoted by a unique index and the concatenation is written right-to-left as is usual for matrix multiplication. Sometimes the vector products offer simplifications that can reduce the complexity of the overall motion.

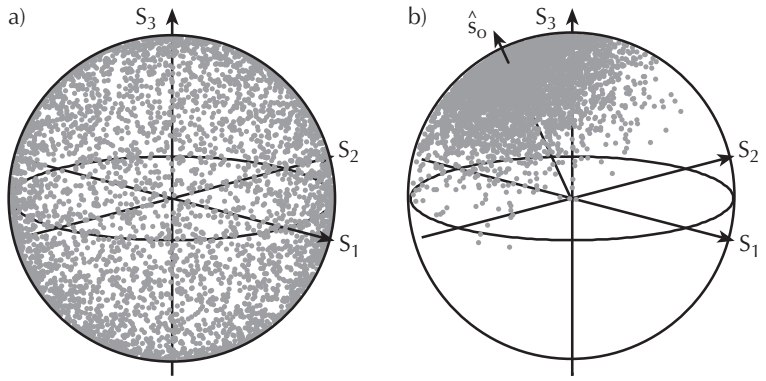


**Fig. 2.8.** Uniform and biased scattering through operator $R$. a) Uniform scattering. $\hat{r}$ points in any direction with equal likelihood. $\varphi$ is uniformly distributed. b) Biased scattering, $a = 0.05$. $\tilde{R}$ is constructed along $s_3$ and oriented toward $\hat{s}_o$.

As another application, the shortest distance between points $\hat{s}$ and $\hat{t}$ on a unit sphere is along a great circle. The axis normal to the great circle is evidently

$$\hat{r} = \frac{\hat{s} \times \hat{t}}{|\hat{s} \times \hat{t}|} \tag{2.6.42}$$

and the rotation angle between the two points is

$$\cos \varphi_{s-t} = \hat{s} \cdot \hat{t} \tag{2.6.43}$$

Figure 2.7(b) illustrates the motion. $\hat{r}$ is derived from the cross of $\hat{s}$ and $\hat{t}$. Angle $\varphi_{s-t}$ rotates $\hat{s}$ through to $\hat{t}$.

Uniform and biased polarization scattering is useful in connection with polarization-mode dispersion fiber-modelling calculations. The scattering process occurs between any two adjacent birefringent sections and it intended to model the relative alignments of the respective birefringent axes. A uniform scattering process sends the polarization state at the output of one section, $\hat{s}_o$, to any point on the Poincaré sphere with equal probability. That state, $\hat{s}_i$, is then input to the next birefringent section. The biased scattering process weights the scattering along a predetermined direction, often the direction of $\hat{s}_o$. For either uniform or biased scattering an operator $R$ needs to be constructed.

There are two variables contained in $R$, (2.6.25): pointing direction $\hat{r}$ and precession angle $\varphi$. Direction $\hat{r}$ itself has two independent variables, the polar angles of declination and azimuth. Combined, $R$ has three independent variables. The random process is derived using the unit deviate $\tilde{u}$. To have $\hat{r}$ point in any direction on the unit sphere with equal likelihood, the azimuth angle $\phi$ and position along the $s_3$ axis are both uniformly distributed. Also, the precession angle is uniformly distributed to generate precessions with equal

likelihood. The random variable expressions are

$$\tilde{r}_3 = 2\tilde{u} - 1 \tag{2.6.44a}$$

$$\tilde{\phi} = (2\tilde{u} - 1)\,\pi \tag{2.6.44b}$$

$$\tilde{\varphi} = (2\tilde{u} - 1)\,\pi \tag{2.6.44c}$$

Relating $\tilde{r}_3$ to the polar angle as $\tilde{r}_3 = \cos\theta$, the remaining coordinates are

$$\tilde{r}_1 = \sin\theta\cos\tilde{\phi} \tag{2.6.45a}$$

$$\tilde{r}_2 = \sin\theta\sin\tilde{\phi} \tag{2.6.45b}$$

The random variable $\tilde{R}$ can now be constructed. Since uniform scattering is completely symmetric on the unit sphere, the pointing direction $\tilde{r}$ does not need to be oriented toward $\hat{s}_o$. Figure 2.8(a) illustrates an output state scattered on the Poincaré sphere.

Biased scattering is used to enhance the likelihood of rare events, rare events being those where multiple birefringent axes are preferentially aligned, misaligned, or some other construction. To preferentially align the birefringent sections, $\hat{r}$ should be biased to point toward $\hat{s}_o$. A simple way to generate the bias is first to bias $\tilde{r}$ towards the $s_3$ direction using the following formula:

$$\tilde{r}_3 = 2\tilde{u}^{1/a} - 1 \tag{2.6.46}$$

For $a \leq 1$ the bias is toward $+s_3$ and for $a \geq 1$ the bias is toward $-s_3$. The scattering operator $R$ is now biased toward $s_3$ and is denoted $R_3$. Before $R_3$ can be applied to $\hat{s}_o$ the former needs to be rotated into the latter. Following the previous example of the shortest distance between two points in Stokes space, a deterministic operator $R_{3-so}$ is constructed to perform the required rotation. Operator $R_{3-so}$ needs to be calculated only once. Figure 2.8(b) illustrates an output state scattered on the Poincaré sphere and biased toward $\hat{s}_o$.

**Table 2.2.** Jones and Stokes Equivalent Expressions

| **Jones expressions** | **Stokes expressions** |
|---|---|
| $\|s\rangle = \begin{pmatrix} s_x \mathrm{e}^{j\phi_x} \\ s_y \mathrm{e}^{j\phi_y} \end{pmatrix}$ | $\hat{s} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}$ |
| $\|t\rangle = T\,\|s\rangle$ | $\hat{t} = R\,\hat{s}$ |
| $\|t\rangle = T_2 T_1\,\|s\rangle$ | $\hat{t} = R_2 R_1 \hat{s}$ |
| $TT^\dagger = T^\dagger T = I$ | $RR^\dagger = R^\dagger R = I$ |
| $\|s\rangle = \hat{s}\cdot\vec{\sigma}\,\|s\rangle$ | $\hat{s} = \langle s\,\|\vec{\sigma}\|\,s\rangle$ |
| $\|s\rangle\langle s\| = \frac{1}{2}\,(I + \hat{s}\cdot\vec{\sigma})$ | $\hat{s} = \frac{1}{2}\,\mathrm{Tr}\,(\|s\rangle\langle s\|\,\vec{\sigma})$ |
| $U\vec{\sigma}U^\dagger$ | $R\,\vec{\sigma}$ |
| $U = I\cos(\varphi/2) - j(\hat{r}\cdot\vec{\sigma})\sin(\varphi/2)$ | $R = (\hat{r}\hat{r}\cdot) + \sin\varphi(\hat{r}\times) - \cos\varphi(\hat{r}\times)(\hat{r}\times)$ |
| $U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}$ | $R_{j,k} = \frac{1}{2}\mathrm{Tr}\left(U\sigma_k U^\dagger \sigma_j\right)$ |
| $U_1 = \begin{pmatrix} \mathrm{e}^{j\alpha} & \\ & \mathrm{e}^{-j\alpha} \end{pmatrix}$ | $R_1 = \begin{pmatrix} 1 & & \\ & \cos 2\alpha & \sin 2\alpha \\ & -\sin 2\alpha & \cos 2\alpha \end{pmatrix}$ |
| $U_2 = \begin{pmatrix} \cos\kappa & -j\sin\kappa \\ -j\sin\kappa & \cos\kappa \end{pmatrix}$ | $R_2 = \begin{pmatrix} \cos 2\kappa & & \sin 2\kappa \\ & 1 & \\ -\sin 2\kappa & & \cos 2\kappa \end{pmatrix}$ |
| $U_3 = \begin{pmatrix} \cos\kappa & -\sin\kappa \\ \sin\kappa & \cos\kappa \end{pmatrix}$ | $R_3 = \begin{pmatrix} \cos 2\kappa & -\sin 2\kappa & \\ \sin 2\kappa & \cos 2\kappa & \\ & & 1 \end{pmatrix}$ |
| $\frac{d\,\|s\rangle}{d\varphi} = -j/2(\hat{r}\cdot\vec{\sigma})\,\|s\rangle$ | $\frac{d\hat{s}}{d\varphi} = \hat{r}\times\hat{s}$ |

**Table 2.3.** Spin-Vector Expressions

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ; \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \sigma_3 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}$$

$$\vec{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix}$$

$$a_0 I + \vec{a} \cdot \vec{\sigma} = \begin{pmatrix} a_0 + a_1 & a_2 - j a_3 \\ a_2 + j a_3 & a_0 - a_1 \end{pmatrix}$$

$$R = (\hat{r}\hat{r}\cdot) + \sin\varphi(\hat{r}\times) + \cos\varphi(I - (\hat{r}\hat{r}\cdot))$$

$$\hat{r}\hat{r}\cdot = \begin{pmatrix} r_1 r_1 & r_1 r_2 & r_1 r_3 \\ r_2 r_1 & r_2 r_2 & r_2 r_3 \\ r_3 r_1 & r_3 r_2 & r_3 r_3 \end{pmatrix}, \quad \hat{r}\times = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix}$$

$$H = \exp(\alpha_0/2)\exp(\vec{\alpha}\cdot\vec{\sigma}/2) = \mathrm{e}^{\alpha_0/2}\ \left[\ I\cosh(\alpha/2) + (\hat{\alpha}\cdot\vec{\sigma})\sinh(\alpha/2)\ \right]$$

$$T = \exp(-j\,\beta_0/2)\exp\left(-j\vec{\beta}\cdot\vec{\sigma}/2\right) = \mathrm{e}^{-j\,\beta_0/2}\ \left[\ I\cos(\beta/2) - j(\hat{\beta}\cdot\vec{\sigma})\sin(\beta/2)\ \right]$$

## References

1. O. Aso, I. Ohshima, and H. Ogoshi, "Unitary-conserving construction of the Jones matrix and its applications to polarization-mode dispersion analysis," *Journal of the Optical Society of America A*, vol. 14, no. 8, pp. 1988–2005, Aug. 1997.
2. D. M. Brink and G. R. Satchler, *Angular Momentum*, 3rd ed. Oxford: Oxford Science Publications, 1999.
3. N. Frigo, "A generalized geometric representation of coupled mode theory," *IEEE Journal of Quantum Electronics*, vol. QE-22, no. 11, pp. 2131–2140, 1986.
4. N. Gisin and B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers," *Optics Communications*, vol. 142, pp. 119–125, Oct. 1997.
5. J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000. [Online]. Available: http://www.pnas.org
6. M. Rose, *Elementary Theory of Angular Momentum*. New York: Dover Publications, 1995.
7. J. J. Sakurai, *Modern Quantum Mechanics*. New York: Addison–Wesley, 1985.
8. G. Strang, *Linear Algebra and its Applications*, 3rd ed. New York: Harcourt Brace Jovanovich College Publishers, 1988.

http://www.springer.com/978-0-387-22493-0

Polarization Optics in Telecommunications
Damask, J.N.
2005, XVI, 522 p. 202 illus., Hardcover
ISBN: 978-0-387-22493-0