



NRL/MR/5650--13-9503

Modeling Interferometric Structures with Birefringent Elements: a Linear Vector-Space Formalism

NICHOLAS J. FRIGO

*U.S. Naval Academy
Physics Department
Annapolis, Maryland*

VINCENT J. URICK

FRANK BUCHOLTZ

*Photonics Technology Branch
Optical Sciences Division*

November 12, 2013

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 12-11-2013		2. REPORT TYPE Memorandum		3. DATES COVERED (From - To) 06-01-2013 – 08-12-2013	
4. TITLE AND SUBTITLE Modeling Interferometric Structures with Birefringent Elements: a Linear Vector-Space Formalism				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 62271N	
6. AUTHOR(S) Nicholas J. Frigo, ¹ Vincent J. Urick, and Frank Bucholtz				5d. PROJECT NUMBER	
				5e. TASK NUMBER EW-271-003	
				5f. WORK UNIT NUMBER 6582	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Research Laboratory, Code 5650 4555 Overlook Avenue, SW Washington, DC 20375-5320				8. PERFORMING ORGANIZATION REPORT NUMBER NRL/MR/5650--13-9503	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research One Liberty Center 875 North Randolph Street, Suite 1425 Arlington, VA 22203-1995				10. SPONSOR / MONITOR'S ACRONYM(S) ONR	
				11. SPONSOR / MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES ¹ U.S. Naval Academy, Physics Department, Annapolis, MD					
14. ABSTRACT We outline a formalism for modeling structures containing birefringent elements and directional couplers. The traditional description of couplers (as single polarization devices with two input and output ports) and fibers (as single port devices with two polarizations) is generalized into a four-dimensional vector space with convenient coordinate transformations between the traditional views. This enables system modelers to introduce complex elements and scenarios under a single formalism, switching between the two traditional views with a single 4×4 matrix transformation. We generalize an earlier formulation to include extra degrees of freedom in rotation, coupling, and birefringence matrix operators. The technique is applied to a microwave photonics setting and a quantitative description is developed that illustrates how normal mode loss in the interferometer's couplers leads to a polarization-independent phase offset in the output arms.					
15. SUBJECT TERMS Coupled mode theory Fiber optics Birefringence Polarization Fiber-optic coupler Interferometer					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Unclassified Unlimited	18. NUMBER OF PAGES 29	19a. NAME OF RESPONSIBLE PERSON Vincent J. Urick
a. REPORT Unclassified Unlimited	b. ABSTRACT Unclassified Unlimited	c. THIS PAGE Unclassified Unlimited			19b. TELEPHONE NUMBER (include area code) (202) 767-9352

CONTENTS

I	EXECUTIVE SUMMARY	E-1
II	INTRODUCTION	1
III	DESCRIPTION OF FORMALISM	2
	Overview	2
	Waveguide View	4
	Coupler View	4
	Co-ordinate Transformation	5
	Interferometric structure example	6
IV	FORM OF OPERATORS	7
	Polarization dependent loss	7
	Birefringent waveguides	8
	Coupler	9
V	APPLICATION	10
	Generalizing earlier work	10
	Interferometric demodulators in microwave photonic links	13
VI	SUMMARY	17
VII	REFERENCES	18
VIII	APPENDIX-A: ELEMENTS OF COUPLED MODE THEORY IN TWO- AND THREE-DIMENSIONS	19
	2D equation of motion	19
	Suppressing overall phase	19
	Pauli matrices	20
	General 2D solution	20
	Geometric representation	21
	Generalization to vectors with overall phase	23
	Generalization to operators with overall phase	24
	Application to birefringent propagation and optical coupling	26
	Conclusion	26

I EXECUTIVE SUMMARY

We outline a formalism for modeling structures containing birefringent elements and directional couplers. The traditional description of couplers (as single polarization devices with two input and output ports) and fibers (as single port devices with two polarizations) is generalized into a four-dimensional vector space with convenient co-ordinate transformations between the traditional views. This enables system modelers to introduce complex elements and scenarios under a single formalism, switching between the two traditional views with a single 4×4 matrix transformation. We generalize an earlier formulation to include extra degrees of freedom in rotation, coupling, and birefringence matrix operators. The technique is applied to a microwave photonics setting and a quantitative description is developed that illustrates how normal mode loss in the interferometer's couplers leads to a polarization-independent phase offset in the output arms.

II INTRODUCTION

Optical directional couplers enabled advances in fiber optic sensor technology[1], including interferometric demodulation in microwave photonic links[2]. The development of integrated[3] Mach-Zehnder interferometers (MZI) has also enabled dramatic advances in digital optical communications, which increasingly use phase modulation to avoid optical impairments[4, 5]. Evolving system metrics will inevitably tighten component and subsystem requirements, driving more research, experimentation, and development to meet stricter specifications, including detailed system dependence on optical polarization. However, modeling polarization dependence can be complicated by the fact that optical couplers, and the waveguides attached to them, have different natural descriptions.

Consider the Mach-Zehnder interferometer (MZI) illustrated schematically in Fig. 1. Waveguides F_1 and F_2 carry optical signals to coupler C_1 . The natural descriptions for each of the fibers are 2×1 Jones vectors which describe polarization evolution. We call this the “waveguide view:” the waveguides are considered independent, but their polarizations may couple during propagation. On the other hand, directional couplers are usually treated in the scalar “coupler view:” the waveguides are viewed as coupled, but only a single polarization is considered. Since the waveguide view (coupled polarizations, but independent waveguides) and the coupler view (coupled waveguides, but independent polarizations) are fundamentally different, polarization in couplers is often ignored as a practical matter.

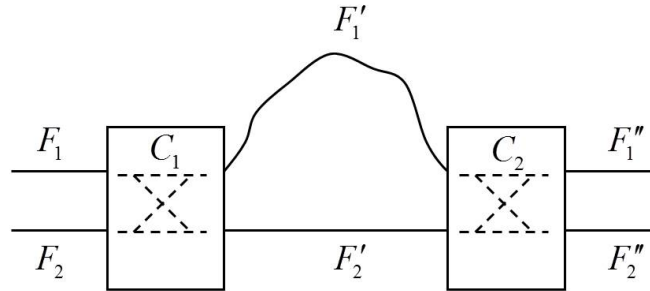


Fig. 1: Basic Mach-Zehnder interferometer. Fibers and couplers have different natural descriptions of the interactions between fields.

Treatments of polarization dependence in couplers are generally based in the component technology: for example, a rigorous development by Chen and Burns[6] treated two birefringent waveguides having arbitrary orientation and twist, with coupling between similar polarizations. Such solutions, which integrate a 4×4 differential equation, are more suitable to device research than quick system modeling. In another approach, more amenable to system modeling, the “fiber view” was used exclusively[7]: the waveguides of Fig. 1 were tracked throughout the structure, including through the couplers. The couplers were modeled as 4×4 matrices with two degrees of freedom (DOF), one DOF for each polarization. The 4×4 matrices they proposed could be considered as a generalization of a Jones matrix[8], with all the attendant modeling

advantages.

This work generalizes that recent paper in several ways. We cast the problem in a 4-dimensional space, as in [7], but we also introduce a simple co-ordinate transformation between the fiber view and the coupler view that permits switching back and forth between the views as needed. This permits symbolic reduction of the 4×4 matrices to 2×2 blocks and allows easier extensions in both the fiber and coupler views. We show later that this permits an easy extension of the model to normal mode loss, for instance. Second, we generalize the expressions for the birefringent waveguide sections to cover two degrees of freedom (DOF) for the birefringence eigenstates, rather than one. In a similar vein, we generalize the coupler expressions to three DOF instead of one, so that modeling of extremely broadband devices can be implemented, for example. In this vein, we generalize the rotation operator, as well. We also describe the formalism using a geometrical representation, so that the comparisons and differences are more readily viewed.

III DESCRIPTION OF FORMALISM

Overview

Our basic approach is illustrated in Fig. 2, upper figure (a) of which is an exploded view of a coupler

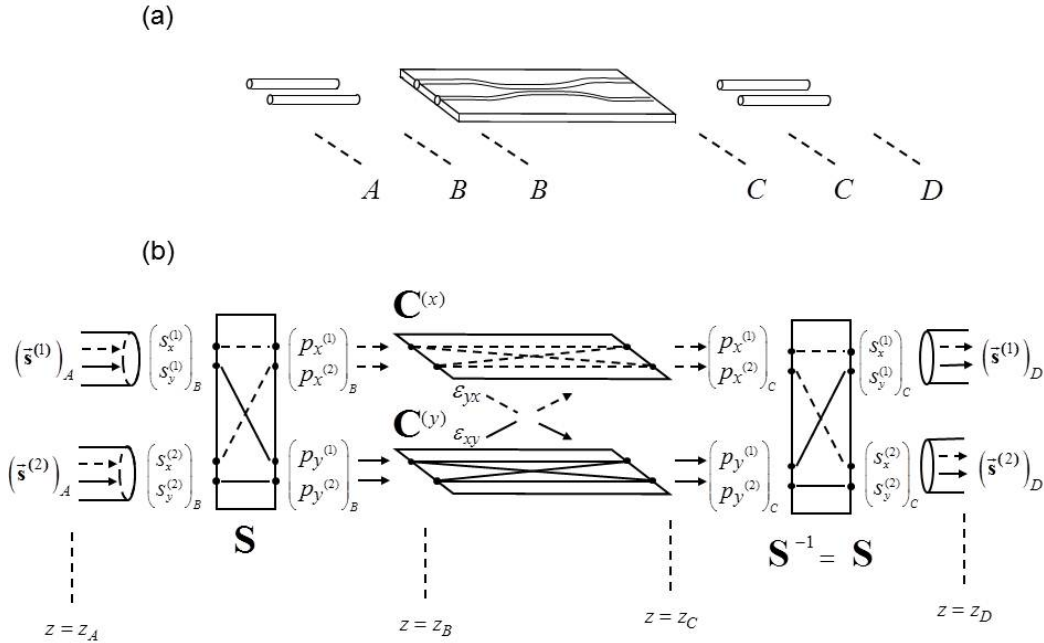


Fig. 2: State space approach. (a) Physical view of input fibers, coupler, and output fibers (b) Exploded view showing ‘x’ (dashed) and ‘y’ (solid) polarizations for the input and output waveguides (1,2) and for the coupler, which is viewed conceptually as two 2×2 couplers, one for each polarization. Transformation matrix ‘ S ’ changes fields from the waveguide view to the coupler view.

with pigtailed, while lower figure (b) is a detailed view of the waveguides, couplers, and their fields. In the

fibers and in the couplers, the x and y polarizations are indicated by dashed and solid lines, respectively.

The physical coupler is viewed conceptually as two couplers, $C^{(x)}$ and $C^{(y)}$, shown in Fig. 2b: there is one for each polarization, and each acts as a conventional 2×2 coupler. Cross-coupling between these two couplers is shown as arrows which change their polarization characteristic in connecting between the two planes. In matrix form, these are represented by 2×2 matrices ε_{xy} and ε_{yx} .

The coupler formulation is different than the waveguide formulation. In the waveguide view, a two dimensional state vector represents the two components of that waveguide's polarization, so that the waveguides' fields are given by

$$\vec{s}^{(1)} = \begin{bmatrix} s_x^{(1)} \\ s_y^{(1)} \end{bmatrix} \quad \vec{s}^{(2)} = \begin{bmatrix} s_x^{(2)} \\ s_y^{(2)} \end{bmatrix}.$$

The joint state of the fields leaving both waveguides and arriving at the coupler, $z = z_B$ in Fig. 2, can be represented by stacking their independent 2-dimensional Jones vectors to form a 4-dimensional vector in the *waveguide view* as

$$\begin{bmatrix} \vec{s}^{(1)} \\ \vec{s}^{(2)} \end{bmatrix}_B = \begin{bmatrix} s_x^{(1)}(z_B) \\ s_y^{(1)}(z_B) \\ s_x^{(2)}(z_B) \\ s_y^{(2)}(z_B) \end{bmatrix}. \quad (1)$$

On the other hand, the fields in the coupler view are represented by two 2-dimensional “port” vectors, one for each coupling plane in Fig 2. Thus, for example, the inputs to the “ x polarization” coupler are given as the amplitudes at the input, z_B , for the two ports:

$$\vec{p}^{(x)}(z_B) = \begin{bmatrix} p_x^{(1)}(z_B) \\ p_x^{(2)}(z_B) \end{bmatrix}.$$

As above, this notation is meant to signify that the fields are described in the coupler view, with the superscript referring to the particular polarization. We alert the reader to the notation.¹ Again, as in Eqn (1), we construct the 4-dimensional vector comprised of these two independent 2-dimensional vectors in the *coupler view* as

$$\begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_B = \begin{bmatrix} p_x^{(1)}(z_B) \\ p_x^{(2)}(z_B) \\ p_y^{(1)}(z_B) \\ p_y^{(2)}(z_B) \end{bmatrix}. \quad (2)$$

But the port amplitudes at interfaces B and C are exactly the same as the waveguide amplitudes there: both descriptions form bases that span the 4-dimensional space describing the fields. Indeed, it must be true that

$$p_x^{(1)} = s_x^{(1)}; p_x^{(2)} = s_x^{(2)}; p_y^{(1)} = s_y^{(1)}; p_y^{(2)} = s_y^{(2)}.$$

In other words, the amplitudes are the same, but they are rearranged when we change the basis for the 4-dimensional space (assumed complete²) from the waveguide view to the coupler view. This rearrangement operation is a co-ordinate transformation in the 4-dimensional space and is symbolized by the operator $\mathbf{S}$ as in Fig 2. After traversing the coupler, the amplitudes are again rearranged from the coupler view to the

¹In the *coupler view*, the *vector's* superscript on the left is the polarization, while the *components'* superscripts in the bracket on the right are the port numbers. These conventions make the matrix algebra and the transformation, described later, more consistent.

²As in [7], we treat losses independently with distinct polarization-dependent elements.

waveguide view by $\mathbf{S}^{-1}$, whereupon their evolution is further described by the Jones matrices for those waveguides.

In the remainder of this section, we will look in more detail at the waveguide view, the coupler view, and the co-ordinate transformation.

Waveguide View

Referring to Fig. 2, polarization states in waveguides 1 and 2 evolve from point A to B , the input to the coupler. From the left, in the waveguide view, propagation from A to B in each fiber is described as a Jones vector evolving under the influence of a 2×2 Jones matrix. Thus, on each of these paths the Jones vectors, $\vec{s}^{(1)}$ and $\vec{s}^{(2)}$, evolve under the influence of their individual Jones matrices (i.e. linear, but not necessarily unitary) for their respective waveguides. There is no coupling between the fibers, so the off-diagonal blocks are zero and we can view the entire process in 2-dimensional blocks:

$$\begin{bmatrix} \vec{s}^{(1)} \\ \vec{s}^{(2)} \end{bmatrix}_B = \begin{bmatrix} s_x^{(1)} \\ s_y^{(1)} \\ s_x^{(2)} \\ s_y^{(2)} \end{bmatrix}_B = \begin{bmatrix} \begin{pmatrix} J_{11}^{(1)} & J_{12}^{(1)} \\ J_{21}^{(1)} & J_{22}^{(1)} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} J_{11}^{(2)} & J_{12}^{(2)} \\ J_{21}^{(2)} & J_{22}^{(2)} \end{pmatrix} \end{bmatrix} \begin{bmatrix} s_x^{(1)} \\ s_y^{(1)} \\ s_x^{(2)} \\ s_y^{(2)} \end{bmatrix}_A = \begin{bmatrix} \mathbf{J}^{(1)} & 0 \\ 0 & \mathbf{J}^{(2)} \end{bmatrix} \begin{bmatrix} \vec{s}^{(1)} \\ \vec{s}^{(2)} \end{bmatrix}_A. \quad (3)$$

The absence of coupling between the fibers, evident in the off-diagonal zero block matrices, is reflected in two independent Jones matrix evolutions

$$\vec{s}^{(i)}(z_B) = \mathbf{J}^{(i)} \vec{s}^{(i)}(z_A),$$

for waveguides $i = 1, 2$.

Coupler View

The formulation for a standard 4-port directional coupler analysis has the same mathematical structure as the Jones calculus, but a different interpretation. The 2×1 vectors in the coupler view now represent the coupled scalar fields at the two physical input ports or output ports, and the coupling is described with a 2×2 coupling matrix. The usual practice is to analyze a single polarization, so that the two amplitudes for the x polarization at the output ports (z_C), for example, are linearly related to the port amplitudes at z_B , the input ports by coupling matrix $\mathbf{C}^{(k)}$

$$\vec{p}_C^{(x)} = \begin{bmatrix} p_x^{(1)}(z_C) \\ p_x^{(2)}(z_C) \end{bmatrix} = \begin{bmatrix} C_{11}^{(x)} & C_{12}^{(x)} \\ C_{21}^{(x)} & C_{22}^{(x)} \end{bmatrix} \begin{bmatrix} p_x^{(1)}(z_B) \\ p_x^{(2)}(z_B) \end{bmatrix} = \mathbf{C}^{(x)} \begin{bmatrix} p_x^{(1)}(z_B) \\ p_x^{(2)}(z_B) \end{bmatrix} = \mathbf{C}^{(x)} \vec{p}_B^{(x)},$$

with a similar equation for the y polarization, and where we have introduced, as above, the 2×2 block coupling matrix form with bold font. If we admit the possibility of coupling between the polarizations, in 2-D block form the coupler is described, similarly to Eqn (3), as

$$\begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_C = \begin{bmatrix} \mathbf{C}^{(x)} & \boldsymbol{\varepsilon}_{xy} \\ \boldsymbol{\varepsilon}_{yx} & \mathbf{C}^{(y)} \end{bmatrix} \begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_B \quad (4)$$

where, for example, the 2×2 matrix $\boldsymbol{\varepsilon}_{xy}$ couples light from the y polarization at input z_B to light in the x polarization at output z_C .

We argue that the $\boldsymbol{\varepsilon}_{ij}$ matrices should be small for several reasons. First, the overlap integral[13] should be small. In the case of plane waves, it would be identically zero for the orthogonal x and y polarizations, and

in adjacent waveguides it should be small and dominated by minor field distortions. Second, the construction of the coupler should force a natural x and y basis by virtue of the planar physical symmetry of the device imposed by its substrate. Any birefringence is likely to be in this basis set, and thus the phase matching between the two x (or y) polarizations should be more efficient than the phase matching between the x and y polarizations. That is, each polarization should couple more strongly to its adjacent guide for similar polarizations than for orthogonal polarizations. This approximation was made explicitly in earlier work[6]. Since both of these effects (i.e., very weak coupling and inefficient phase matching) suggest minimal power transfer, their combination leads us to assume negligible coupling for our basic model. That is, although the model does not require the ε to be zero matrices, we assume, as earlier [7], that they can be ignored and we illustrate the model with that assumption.

In matrix form, then, in traversing the coupler from z_B to z_C , the fields can be described in the coupler view as

$$\begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_C = \begin{bmatrix} p_x^{(1)} \\ p_x^{(2)} \\ p_y^{(1)} \\ p_y^{(2)} \end{bmatrix}_C = \begin{bmatrix} C_{11}^{(x)} & C_{12}^{(x)} & 0 & 0 \\ C_{21}^{(x)} & C_{22}^{(x)} & 0 & 0 \\ 0 & 0 & C_{11}^{(y)} & C_{12}^{(y)} \\ 0 & 0 & C_{21}^{(y)} & C_{22}^{(y)} \end{bmatrix} \begin{bmatrix} p_x^{(1)} \\ p_x^{(2)} \\ p_y^{(1)} \\ p_y^{(2)} \end{bmatrix}_B = \begin{bmatrix} \mathbf{C}^{(x)} & 0 \\ 0 & \mathbf{C}^{(y)} \end{bmatrix} \begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_B.$$

As in the case for the fiber view, the off-diagonal zero block matrices indicate the lack of polarization coupling in the coupler, and thus there are two independent 2×2 block coupling matrices on the diagonal for the two polarizations

$$\vec{p}^{(i)}(z_C) = \mathbf{C}^{(i)} \vec{p}^{(i)}(z_B),$$

for polarizations $i = x, y$.

Co-ordinate Transformation

The block diagonal forms for the waveguide and coupler regions would naively suggest that the interferometer of Fig. 1 could be treated as a product of block diagonal matrices, which would lead to two separable two dimensional vectors. However, the $\mathbf{J}$ and $\mathbf{C}$ matrices have different bases as discussed above. Thus, a co-ordinate transformation is needed to move between the two physical descriptions. In light of the above equation, one sees that the co-ordinate transformation taking the waveguide view to the coupler view is

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{S}^T = \mathbf{S}^{-1} : \quad (5)$$

in other words, a simple transposition of the 2nd and 3rd rows (or columns). Thus

$$\begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_B = \begin{bmatrix} p_x^{(1)}(z_B) \\ p_x^{(2)}(z_B) \\ p_y^{(1)}(z_B) \\ p_y^{(2)}(z_B) \end{bmatrix} = \begin{bmatrix} s_x^{(1)}(z_B) \\ s_x^{(2)}(z_B) \\ s_y^{(1)}(z_B) \\ s_y^{(2)}(z_B) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x^{(1)}(z_B) \\ s_y^{(1)}(z_B) \\ s_x^{(2)}(z_B) \\ s_y^{(2)}(z_B) \end{bmatrix} = \mathbf{S} \begin{bmatrix} \vec{s}^{(1)} \\ \vec{s}^{(2)} \end{bmatrix}_B,$$

where the fields are stacked in the coupler view on the left side of the equation, and stacked in the waveguide view on the right side of the equation.

This formulation shows that at point B, the waveguide/coupler junction, we can transform our viewpoint by making a co-ordinate transformation. That is, we go from the waveguide view, in which our 4-D space consists of two Jones vectors $\vec{s}^{(1)}$ and $\vec{s}^{(2)}$, to the coupler view, in which case the space is spanned by port vectors for the two polarizations, namely $\vec{p}^{(x)}$ and $\vec{p}^{(y)}$. Because $\mathbf{S}^{-1} = \mathbf{S}$, we could also transform the other way:

$$\begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_B = \mathbf{S} \begin{bmatrix} \vec{s}^{(1)} \\ \vec{s}^{(2)} \end{bmatrix}_B \iff \begin{bmatrix} \vec{s}^{(1)} \\ \vec{s}^{(2)} \end{bmatrix}_B = \mathbf{S}^{-1} \begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_B = \mathbf{S} \begin{bmatrix} \vec{p}^{(x)} \\ \vec{p}^{(y)} \end{bmatrix}_B$$

Interferometric structure example

The formalism enables us to write down structures such as those in Fig. 3 by operating either with Jones

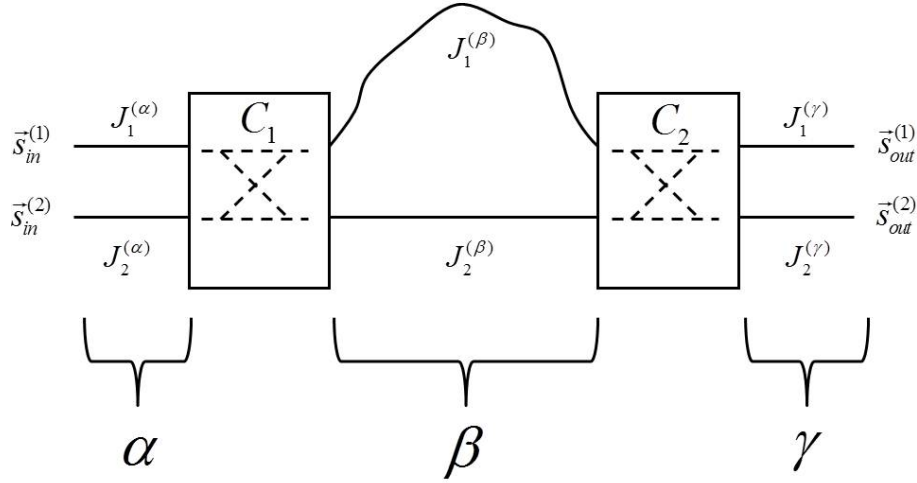


Fig. 3: Interferometric structure. Input fields in two waveguides propagate through sections α , β , and γ , and are coupled in couplers C_1 and C_2 . At the transitions, the co-ordinates are transformed between waveguide and coupler views by matrix $\mathbf{S}$ (see text).

matrices on waveguide sections, or with coupling matrices on couplers, so long as we use $\mathbf{S}$ to transform between the views. Given a structure such as shown in Fig. 3, we can describe the field evolution by stepping through the structure, as described below.

We can relate the input and output Jones vectors in Fig. 3 by inspection using 2×2 matrices (with 2×2 block matrix elements) for the operators and a 2×1 vector (with 2×1 vectors for elements) for the field vectors:

$$\begin{bmatrix} \vec{s}_{out}^{(1)} \\ \vec{s}_{out}^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_1^{(\gamma)} & 0 \\ 0 & \mathbf{J}_2^{(\gamma)} \end{bmatrix} \left(\mathbf{S} \begin{bmatrix} \mathbf{C}_2^{(x)} & 0 \\ 0 & \mathbf{C}_2^{(y)} \end{bmatrix} \mathbf{S} \right) \begin{bmatrix} \mathbf{J}_1^{(\beta)} & 0 \\ 0 & \mathbf{J}_2^{(\beta)} \end{bmatrix} \left(\mathbf{S} \begin{bmatrix} \mathbf{C}_1^{(x)} & 0 \\ 0 & \mathbf{C}_1^{(y)} \end{bmatrix} \mathbf{S} \right) \begin{bmatrix} \mathbf{J}_1^{(\alpha)} & 0 \\ 0 & \mathbf{J}_2^{(\alpha)} \end{bmatrix} \begin{bmatrix} \vec{s}_{in}^{(1)} \\ \vec{s}_{in}^{(2)} \end{bmatrix}, \quad (6)$$

where we group the $\mathbf{SCS}$ matrices in parentheses to connect to the earlier work[7]. That is, those operators transform from the waveguide view to the coupler view, apply the coupler operator to the state in coupler view, then transform back to the waveguide view: this is a similarity transform of $\mathbf{C}$ into the waveguide view used earlier[7].

In words, we describe this equation (from right to left) as

- transporting the waveguide's fields across the α waveguides;
- changing to the coupler view, operating with the first coupler, and then changing back to the waveguide view;
- transporting the waveguide's fields across the β waveguides;
- changing to the coupler view, operating with the second coupler, and then changing back to the waveguide view;
- transporting the waveguide's fields across the γ waveguides.

One can analyze more complicated structures by walking through them in a similar way, using the 2×2 block form. We note a distinct advantage to this structure: by chaining the operators to keep them in 2×2 block form, we may extend operators easily at each stage. That is, for instance, if we wanted to change the form of $C_2^{(x)}$ to model normal mode losses, it is possible to do so directly by modifying $C_2^{(x)}$'s matrix in the coupler view, where it is a 2×2 matrix.

IV FORM OF OPERATORS

In this section, we summarize the forms that the 2×2 operators take in expressions such as Eqn (6), and compare our expressions to those of Ref [7].

Polarization dependent loss

The simplest optical instrument is a polarizer, as in Fig. 4, illustratively described in the waveguide view as

$$\mathbf{P}(0) = \begin{bmatrix} p_x & 0 \\ 0 & p_y \end{bmatrix} \quad 0 \leq p_x, p_y \leq 1$$

where the diagonal terms describe the field transmissions in the x and y polarizations [8]: an ideal x polarizer would have $p_x = 1$, $p_y = 0$. The rotation matrix

$$\mathbf{R}(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}, \quad (7)$$

which actively rotates a 2D vector through angle α , can be used to describe a polarizer oriented at an angle. The matrix product

$$\mathbf{P}(\alpha) = \mathbf{R}(\alpha) \mathbf{P}(0) \mathbf{R}^{-1}(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} p_x & 0 \\ 0 & p_y \end{bmatrix} \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \quad (8)$$

describes a polarizer whose x axis is oriented as shown in Fig. 4: the components of the field are first rotated by $-\alpha$, then operated on by $\mathbf{P}(0)$ expressed in its own basis, followed by a rotation of $+\alpha$ to restore the original orientation. This form agrees with Eqn (5) of the earlier formulation[7], where we have corrected a typographical error, and their first matrix of Eqn (6). We will generalize this below.

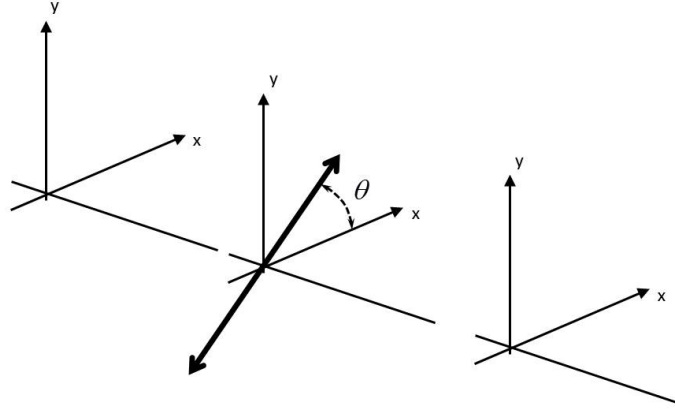


Fig. 4: Polarizer in waveguide view. Polarizer is oriented at angle θ with respect to the $x - y$ fiber bases. Its operator is a similarity transformation of the operator in the polarizer's basis.

Birefringent waveguides

Propagation in birefringent waveguides is described naturally with the Jones calculus[8]-[10], and its extension to include overall optical phase [11, 12], reviewed in Appendix A (Section VIII). Lossless propagation is modeled with the general unitary matrix

$$\mathbf{J}(\bar{\phi}, \phi, \hat{\beta}_0) = e^{-j(\bar{\phi} + (\phi/2)\hat{\beta}_0 \cdot \vec{\sigma})} = e^{-j\bar{\phi}} \left[\cos \frac{\phi}{2} \sigma_0 - j \sin \frac{\phi}{2} (\hat{\beta}_0 \cdot \vec{\sigma}) \right] \quad (9)$$

or

$$\mathbf{J}(\bar{\phi}, \phi, \hat{\beta}_0) = e^{-j2\bar{\phi}/2} \begin{bmatrix} \cos \frac{\phi}{2} - j \sin \frac{\phi}{2} (\beta_3) & -j \sin \frac{\phi}{2} (\beta_1 - j\beta_2) \\ -j \sin \frac{\phi}{2} (\beta_1 + j\beta_2) & \cos \frac{\phi}{2} - j \sin \frac{\phi}{2} (-\beta_3) \end{bmatrix}. \quad (10)$$

This equation has all four degrees of freedom (DOF) permitted by unitary matrices: one DOF for the overall phase $\bar{\phi}$, shown in Eqn (37) in Section VIII; one DOF for the phase difference, $\phi = 2\bar{\beta}z$, that the operator causes between components in the operator's eigenstates; and two DOF for the unit vector $\hat{\beta}_0$ the location of which represents the operator's eigenstate.

A geometrical representation of these four DOF is shown in Fig. 5, which illustrates $\mathbf{J}$ operating on an arbitrary state vector [11, 12]. Note that, compared to the conventional representation, state $|s_i\rangle$ is represented by the 3D vector $\vec{s}_i$ as well as a “fiducial paddle” which permits an overall phase to be represented. In this representation, the “slow” eigenstate (eigenvalue +1) of $\mathbf{J}$ is represented by the unit vector $\hat{\beta}$, whose components are the $(\beta_1, \beta_2, \beta_3)$ in the matrix representation Eqn (10). Operator $\mathbf{J}$, in the 3D representation, rotates states about $\hat{\beta}$ in the positive (right-hand rule) sense by angle $\phi = 2\bar{\beta}z$, changing the latitude and longitude of the state, and hence the relative amplitudes and phases of the optical components. In addition to changing the relative phases of $|s\rangle$'s components, represented by the changing position of $\vec{s}$, the operator also changes the overall optical phase, which is represented by a fiducial paddle attached to $\vec{s}$. Thus, in rotating through angle ϕ , the orientation of the initial fiducial paddle at $\vec{s}_i$ is changed to the grey shaded paddle at $\vec{s}_f$. The overall phase change is represented by a rotation of the fiducial paddle by an additional

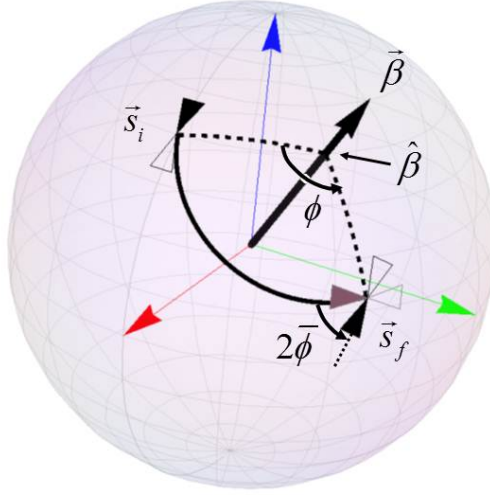


Fig. 5: Geometrical representation of general unitary matrix. Operator $\mathbf{J}$ in Eqns (9,10) with eigenbasis $\hat{\beta}$ rotates state about $\hat{\beta}$ through angle ϕ . Simultaneously, the fiducial paddle representing overall phase, is rotated about $\vec{s}$ by angle $2\bar{\phi}$.

$2\bar{\phi}$ radians in the positive sense. The four DOF for this unitary transformation are thus completely identified in this representation.

Conversely, any unitary matrix, since it must have the form of Eqn (10), can thus be represented by its equivalent $\hat{\beta}$, $\bar{\phi}$, and ϕ , by projecting that matrix onto the basis set $(\sigma_0, \sigma_1, \sigma_2, \sigma_3)$ and identifying the phases [9]-[12]. We will use this to compare our approach to the earlier model [7].

Coupler

Ideal coupling matrices are unitary matrices similar to the $\mathbf{J}$ operators discussed above and reviewed in Appendix A (Section VIII): they conserve power and are linear in the $\vec{E}$ fields. Consequently their $\mathbf{C}$ matrices can be described with the same 4 parameters we used to describe propagation operators for birefringent fibers: what changes is that the basis states are now unit fields in waveguides associated with ports 1 and 2, rather than unit fields in x and y polarizations.

$$\mathbf{C}(\bar{\phi}, \phi, \hat{\kappa}) = e^{-j(\bar{\phi} + (\phi/2)\hat{\kappa} \cdot \vec{\sigma})} = e^{-j\bar{\phi}} \left[\cos \frac{\phi}{2} \sigma_0 - j \sin \frac{\phi}{2} (\hat{\kappa} \cdot \vec{\sigma}) \right] \quad (11)$$

or

$$\mathbf{C}(\bar{\phi}, \phi, \hat{\kappa}) = e^{-j2\bar{\phi}/2} \begin{bmatrix} \cos \frac{\phi}{2} - j \sin \frac{\phi}{2} (\kappa_3) & -j \sin \frac{\phi}{2} (\kappa_1 - j\kappa_2) \\ -j \sin \frac{\phi}{2} (\kappa_1 + j\kappa_2) & \cos \frac{\phi}{2} - j \sin \frac{\phi}{2} (-\kappa_3) \end{bmatrix}. \quad (12)$$

Thus, $2\bar{\phi} = 2\bar{\kappa}z$ describes the overall optical length of the coupler, ϕ describes phase differences between fields in the waveguides, and $\hat{\kappa}$ defines the eigenstates of the coupler in terms of the waveguide modes. A slight difference in the respective interpretations of the two situations is that the coupling components are

conventionally viewed as having a different origin. Components κ_1 and κ_2 are obtained from the overlap integrals between the fields in the two modes [13, 14]: the field from one guide creates a polarization in the other guide, which drives the wave in the first guide, and vice versa. The third component, κ_3 , is associated with the de-tuning of the input field from the ideal phase-matched condition.

V APPLICATION

Generalizing earlier work

Earlier work[7] introduced a formalism using 4×4 matrix operators to describe polarization-dependent interferometers. Their approach used coupling, birefringence, and polarization-dependent loss (PDL) operators: $C(\theta_x, \theta_y)$, $D(\phi)$, and $T(\sigma_{nm})$, respectively. Additionally, a rotation matrix $R(\alpha_a, \alpha_b)$ was introduced that transformed those operators to different PDL and birefringence axes on waveguides a and b , respectively. The operators are cast in what we call the “waveguide view” exclusively. In this subsection we generalize those operators.

Operator for coupler

The coupling operator, again for $i \rightarrow -j$, was given as [7]

$$\mathbf{C}(\theta_x, \theta_y) = \begin{bmatrix} \cos(\theta_x) & 0 & -j \sin(\theta_x) & 0 \\ 0 & \cos(\theta_y) & 0 & -j \sin(\theta_y) \\ -j \sin(\theta_x) & 0 & \cos(\theta_x) & 0 \\ 0 & -j \sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix}, \quad (13)$$

where θ_x and θ_y represent the different splitting ratios for the couplers on the two polarization planes in Fig.2. Since this operator is in the waveguide view, we can find the corresponding operator in the coupler view, $\mathbf{C}_c$, by a similarity transform:

$$\mathbf{C}(\theta_x, \theta_y) = \mathbf{S} \mathbf{C}_c \mathbf{S}^{-1} \implies \mathbf{C}_c = \mathbf{S}^{-1} \mathbf{C}(\theta_x, \theta_y) \mathbf{S}. \quad (14)$$

By virtue of $\mathbf{S}^{-1} = \mathbf{S}$, this means that

$$\mathbf{C}_c = \mathbf{S} \mathbf{C}(\theta_x, \theta_y) \mathbf{S} = \begin{bmatrix} \cos(\theta_x) & -j \sin(\theta_x) & 0 & 0 \\ -j \sin(\theta_x) & \cos(\theta_x) & 0 & 0 \\ 0 & 0 & \cos(\theta_y) & -j \sin(\theta_y) \\ 0 & 0 & -j \sin(\theta_y) & \cos(\theta_y) \end{bmatrix}. \quad (15)$$

This, in turn, means that the coupling matrices for each of the x and y polarizations have the very specific form

$$\mathbf{C}^{(k)}(\theta_k) = \begin{bmatrix} \cos(\theta_k) & -j \sin(\theta_k) \\ -j \sin(\theta_k) & \cos(\theta_k) \end{bmatrix} \implies \mathbf{C}_c = \begin{bmatrix} C^{(x)}(\theta_x) & 0 \\ 0 & C^{(y)}(\theta_y) \end{bmatrix}, \quad (16)$$

where each coupler operator is seen to have one degree of freedom, θ_n . But as we’ve seen (both in Section IV and in Appendix A, Section VIII), the coupling operator has 3 DOF (or 4 if one includes overall phase), so the 2×2 coupling matrices in [7] can be generalized by

$$\mathbf{C}^{(k)}(\theta_k) \longrightarrow \mathbf{C}^{(k)}(\bar{\phi}_k, \phi_k, \hat{\kappa}_k) \implies \mathbf{C}_c = \begin{bmatrix} C^{(x)}(\bar{\phi}_x, \phi_x, \bar{\kappa}_x) & 0 \\ 0 & C^{(y)}(\bar{\phi}_y, \phi_y, \bar{\kappa}_y) \end{bmatrix} \quad (17)$$

in the coupler view. We now illustrate why one may want the more general form.

For an ideal coupler with identical waveguides placed symmetrically about a central plane orthogonal to the substrate plane, general symmetry arguments show that the eigenstates are even and odd functions [15]. The usual approach [13] uses the unperturbed waveguide modes $|1\rangle$ and $|2\rangle$ as the basis set to form the coupler's eigenmodes

$$|\psi_s\rangle = |e\rangle = \frac{1}{\sqrt{2}} (|1\rangle + |2\rangle) \quad |\psi_f\rangle = |o\rangle = \frac{1}{\sqrt{2}} (|1\rangle - |2\rangle) \quad (18)$$

where we have explicitly labeled the even and odd superposition states.

A general power splitter was modeled by introducing an angle that corresponds to differing coupling lengths [7]: reducing the 4 DOF in Eqn (12) to one ($\bar{\phi} = \kappa_2 = \kappa_3 = 0$), the operator $\mathbf{J}$ in Eqn (12) reduces to

$$\mathbf{C}_{ideal}(\theta) = \begin{bmatrix} \cos \theta & -j \sin \theta \\ -j \sin \theta & \cos \theta \end{bmatrix} \xrightarrow{ideal\ 3\ dB} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix}, \quad (19)$$

where we have used the $e^{+j\omega t}$ convention.

We can use the geometric representation to show the difference between two couplers, both of which split power equally, but have radically different forms. In the geometrical representation, $|1\rangle$, all light in guide 1, is represented by $\hat{e}_3$; $|2\rangle$, all light in guide 2, is represented by $-\hat{e}_3$; and $|e\rangle$, the operator's slow eigenstate, is represented by $\hat{\kappa} = \hat{e}_1$, the axis about which the input states evolve. For an ideal 3 dB coupler, an input into port 1 would evolve into equal power in both guides at the output and hence would correspond to a rotation of $\pi/2$ about the $\hat{e}_1$ axis, shown as the first operation in Fig. 6a. If this ideal coupler were followed by another identical coupler to make a zero path length interferometer, the system would evolve as the right side of Fig. 6a. This form (or, equivalently, $\theta \rightarrow -\theta$) is the only form possible for a 3 dB coupler with one DOF.

An extreme example of another 3 dB coupler is given in Fig. 6b. In this case, the coupler is strongly detuned: κ_3 in this coupler (Δ in [13]) has the same magnitude as κ_1 , but it is still a 3 dB coupler in the sense that it would evenly split the power introduced at an input port. The system evolution is quite different: the output is identical to the input³, and no fixed phase shift between the stages of the system in Fig. 6a can replicate the behavior of Fig. 6b. Using Eqn (17) permits greater modeling flexibility.

Operator for propagation in birefringent media

Birefringence is manifested as different propagation phases for different polarizations, and in the normal $x - y$ basis, this is represented with different phase terms on the diagonal. To permit eigenstates other than the basis set, Ref [7] uses the same rotation matrix, Eqn(7), used for PDL in Eqn (8), for each of the 2×2 diagonal blocks as

$$\mathbf{R}(\alpha) \mathbf{D}(\phi_n) \mathbf{R}^{-1}(\alpha) : \quad \mathbf{D}(\phi_n) = \begin{bmatrix} e^{-j\phi_x} & 0 \\ 0 & e^{-j\phi_y} \end{bmatrix} = e^{-j\bar{\phi}} \begin{bmatrix} e^{-j\phi/2} & 0 \\ 0 & e^{+j\phi/2} \end{bmatrix} \quad (20)$$

where we have made the substitution $i \rightarrow -j$ to convert to the $e^{j\omega t}$ convention [10, 11, 12]. $\mathbf{D}$ contains two DOF: common phase $\bar{\phi} = (\phi_x + \phi_y)/2$, and differential phase $\phi = \phi_x - \phi_y$. Performing the multiplication in Eqn (20), we find that

$$\mathbf{R}(\alpha) \mathbf{D}(\phi) \mathbf{R}^{-1}(\alpha) = e^{-j2\bar{\phi}/2} \left[\cos \phi \sigma_0 - j \sin \phi (\hat{\beta} \cdot \vec{\sigma}) \right] : \quad \hat{\beta} \longrightarrow \begin{bmatrix} \sin 2\alpha \\ 0 \\ \cos 2\alpha \end{bmatrix}, \quad (21)$$

³Technically, all states would have picked up an additional π radians of phase shift, as can be seen by tracking the fiducial paddles.

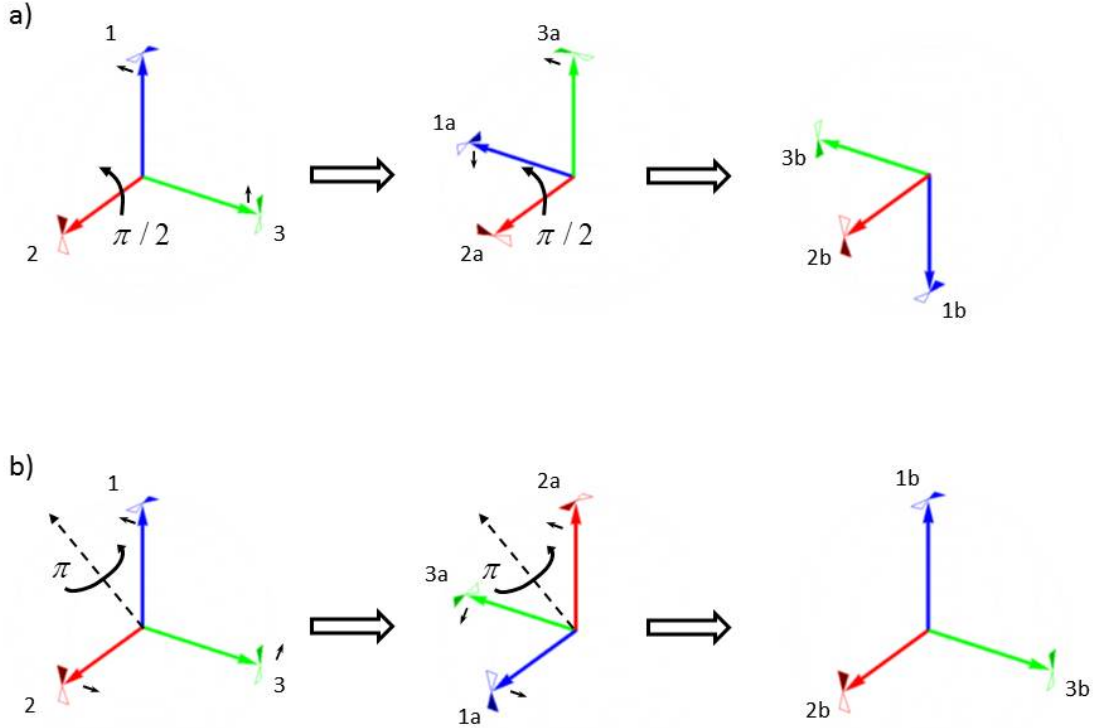


Fig. 6: Two different ‘3 dB’ couplers in an interferometer. (a) Ideal coupler, zero detuning. Each coupler transit is represented by a $\pi/2$ rotation about the even state, $(\hat{\mathbf{e}}_1)$. (b) Detuned coupler. Waveguide 1 has greater index, which would prevent full power transfer. Each coupler transit is represented by a π rotation about $(\hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_3)/\sqrt{2}$

showing that $\hat{\beta}$ has only one degree of freedom, not two. (We might also have noted that, instead of 4 DOF as in Eqn (17), the operator in Eqn (21) has 3 DOF: two for the birefringence $\mathbf{D}(\phi_n)$, and one for α , the elevation of $\hat{\beta}$ above $\mathbf{e}_1$ in the $\mathbf{e}_1 - \mathbf{e}_3$ plane.) The similarity transformation implied in Eqns (20,21) thus corresponds to eigenstates $\hat{\beta}$ limited to linear birefringence.

In contrast, the birefringence operator $\mathbf{J}$ in Eqn (10) has 4 DOF, and has eigenstates with two DOF. In fact, it can be shown that $\mathbf{J}$ is a similarity transform of the birefringence operator $\mathbf{D}(\phi_n)$, but with eigenstates corresponding to the normal form (see Appendix A, Eqns (9,10), and Fig. 5)

$$|\psi\rangle \sim \vec{s} \sim \begin{bmatrix} \cos \frac{\theta}{2} e^{-j\phi/2} \\ \sin \frac{\theta}{2} e^{+j\phi/2} \end{bmatrix}, \quad (22)$$

which expresses two DOF for the slow birefringence eigenstate.

In terms of modeling, the two formulations would thus lead to different statistical distributions in birefringent links and their interference. In principle, the same argument can be made about the representation of the rotated partial polarizer: rotation matrix Eqn (7) allows PDL for linear eigenstates only, while $\mathbf{J}$ permits more general eigenstates.

Summary

We have shown that generalizing the rotation operator opens more degrees of freedom to the coupler, birefringence, and PDL operators in Ref [[7]]. But the basic transformation introduced with Eqn (5), is also useful in extending analyses. For instance, the transformation shows the structure of the differential equation in Ref [6] treats only coupling between similar polarizations. We have made the same assumptions here, but the transformation is useful in comparing different formulations. By using a transformation between the waveguide and coupler views, we can treat all components as simpler block diagonal matrices, at the cost of introducing a similarity transform. In the next subsection, we show that there are advantages to this tradeoff: we apply this advantage to an analysis of demodulators in microwave photonic links.

Interferometric demodulators in microwave photonic links

Overview

In this section, we apply the formalism to address the modeling of interferometers with birefringent elements, schematically shown in Figs. 1 and 3. The starting point is the operator in Eqn (6), which is a 4×4 matrix operator taking the input state to the output state. Generally, one is interested in the detected power, so the inner product of the final state is taken with itself. This operation can be performed either by constructing an overall matrix by pre-multiplication by the matrix's adjoint, or by taking the inner product of the output directly. Application packages such as *Mathematica* make keeping track of the fields more manageable.

General development

To clarify the development of the output fields associated with Eqn (6), we show an exploded diagram such as Fig. 7 to help identify the matrix elements in the analysis. Fig. 7a shows each of the diagonal subsystems, the 2×2 matrices, as a shaded rectangle, while Fig. 7b identifies the functions of the individual matrix elements for both the couplers and the waveguides.

As an illustrative example, assume that the input is on port 1 and has x polarization, i.e. $E_{1x} = 1$ and all other input fields are zero. We analyze the system to find the powers at exit ports 1 and 2, the powers that will be collected by fibers 1 and 2, respectively. Defining the $i - j$ coupling coefficient in Fig. 7 for coupler n and polarization k as $(\mathbf{C}_n^{(k)})_{ij} = c_{nkij}$, the fields at the output are given by

$$\vec{\mathbf{E}} = \begin{bmatrix} E_{1x} \\ E_{1y} \\ E_{2x} \\ E_{2y} \end{bmatrix} = \mathbf{C}_2 \mathbf{S} \mathbf{J} \mathbf{S} \mathbf{C}_1 \mathbf{S} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} c_{2x11} J_{11}^{(1)} c_{1x11} + c_{2x12} J_{11}^{(2)} c_{1x21} \\ c_{2y11} J_{21}^{(1)} c_{1x11} + c_{2y12} J_{21}^{(2)} c_{1x21} \\ c_{2x21} J_{11}^{(1)} c_{1x11} + c_{2x22} J_{11}^{(2)} c_{1x21} \\ c_{2y21} J_{21}^{(1)} c_{1x11} + c_{2y22} J_{21}^{(2)} c_{1x21} \end{bmatrix} \quad (23)$$

Thus, for example, the y component on fiber 2 is the superposition of light following two paths: (i) across the top arm of coupler $\mathbf{C}_1^{(x)}$, coupled into the y polarization through waveguide $\mathbf{J}_1$, and then coupled diagonally across $\mathbf{C}_2^{(y)}$ to port 2; and (ii) coupled diagonally across $\mathbf{C}_1^{(x)}$ to waveguide 2, coupled into the y polarization through waveguide $\mathbf{J}_2$, and across the bottom arm of coupler $\mathbf{C}_2^{(y)}$. This is represented as the bottom element on the right of Eqn (23).

These outputs, as in the earlier development[7], are adequate for CW inputs, but need to be modified slightly for modulated inputs. Specifically, the effects of the interferometric waveguides can not be represented by their phase delay alone, since the field's modulation does not appear instantly at the output. The standard scalar field Fourier analysis, for modulation frequencies much less than the carrier frequency, shows

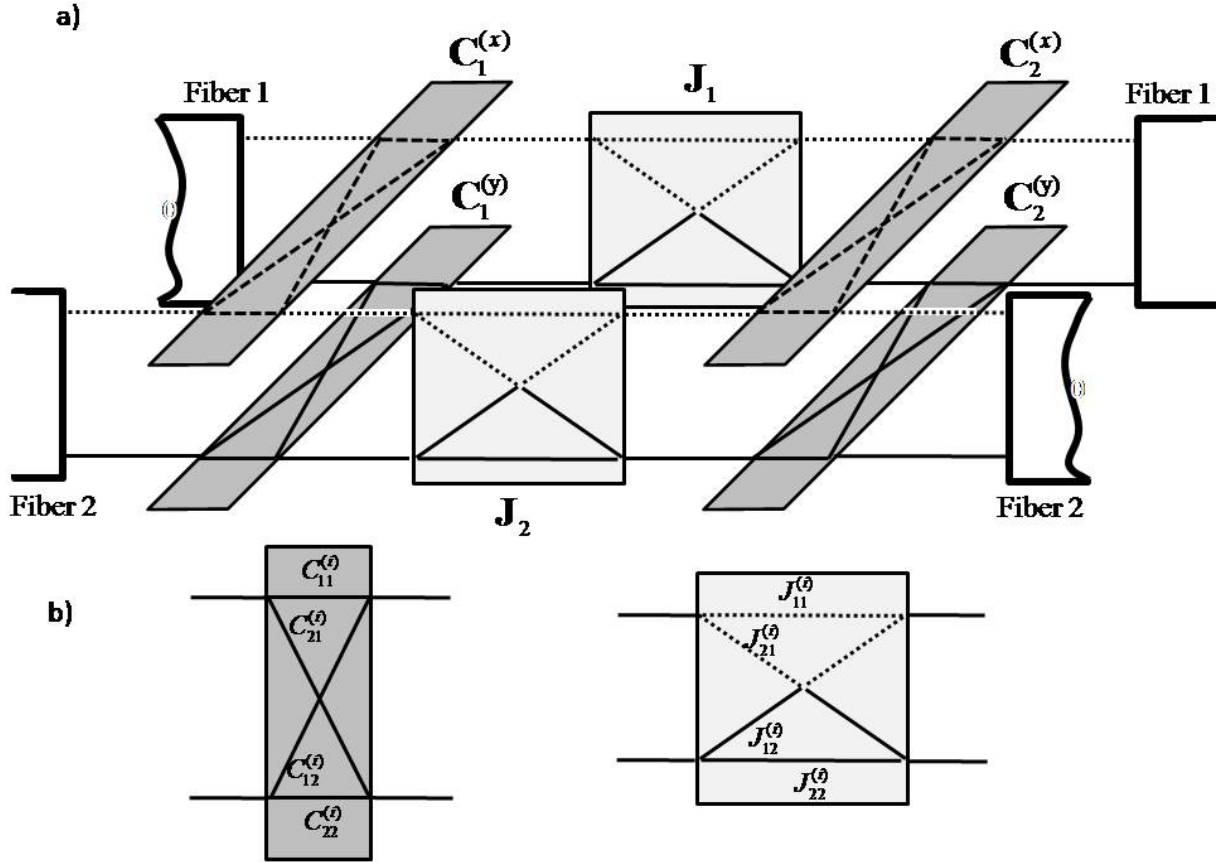


Fig. 7: Interferometer analysis. (a) Schematic as in Fig. 2 showing coupler polarization planes and birefringent fibers. Matrix elements are represented by lines (x are dotted, y are solid) between elements. (b) Definition of matrix elements for coupler plane, C , and waveguides, J .

that the modulation envelope travels at the group velocity, while the carrier travels at the phase velocity[14]. To treat our birefringent waveguides, we assume that polarization mode dispersion (PMD) is unimportant. By this we mean that the modulation envelopes travel at the group velocity through the fiber, but that the differential group delay (DGD) between the principal states of polarization is insignificant for the modulation imposed[10]. In this regime, the modulation envelopes suffer a common group delay τ while the states of polarization of all the spectral components essentially experience the same evolution. We can then treat the fields as in Eqn (23) with the understanding that the modulation envelope must be tracked. For instance,

$$J_{11}^{(1)} E(t) e^{j\omega t} \longrightarrow J_{11}^{(1)} E(t - \tau) e^{j\omega t} \quad (24)$$

where we have subsumed the phase delay into the definition of J , and the $E(t)$ includes both amplitude and phase modulation. Then expansions such as Eqn (23) must be tracked with the fields on which they operate.

Modeling normal mode loss

In this section, we demonstrate the utility of the vector space formalism by modeling couplers with normal mode loss. We exploit vector space techniques to calculate the coefficients in a 2×2 coupler. Given

these matrix elements, one can then put them into the appropriate blocks in Eqn (6). The basic idea is to use another transformation from the usual coupler basis to the normal mode basis for the coupler to model the loss, then transform back to the original basis.

We first note that a symmetric polarization-independent coupler can be viewed in terms of its even and odd normal modes[15] and the matrices which transform co-ordinates between the port co-ordinates and the normal mode co-ordinates.

$$|e\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad |o\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}; \quad S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = S^{-1} = S^\dagger \quad (25)$$

The length of the coupler determines its coupling ratio, and this is determined by the beating between the two modes as they propagate through the length of the coupler. For an ideal coupler, disregarding the common phase, this would give rise to a coupling matrix of

$$C = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-j\theta_C/2} & 0 \\ 0 & e^{+j\theta_C/2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \cos \frac{\theta_C}{2} & -j \sin \frac{\theta_C}{2} \\ -j \sin \frac{\theta_C}{2} & \cos \frac{\theta_C}{2} \end{bmatrix}, \quad (26)$$

namely the ideal generalized coupler of Ref [7]. In terms of the geometrical representation, the coupler is represented by a rotation about the $\hat{e}_1$ through an angle of θ_C : for an ideal $3dB$ coupler, $\theta_C = \pi/2$, and an input to port 1 ($|1\rangle$), would evolve to state $|-\rangle$, for instance.

If both the even and odd modes experience the same loss, the output is merely scaled, so to model differential mode loss, we assume that the E field transmission in the even mode is unity while that in the odd mode is $p \leq 1$. Assuming an otherwise ideal coupler, the coupling matrix would then be

$$C = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-j\pi/4} & 0 \\ 0 & pe^{+j\pi/4} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{e^{-j\pi/4}}{2} \begin{bmatrix} 1+jp & 1-jp \\ 1-jp & 1+jp \end{bmatrix}. \quad (27)$$

For $p = 1$ and input on port 1, one sees that the amplitudes at the output ports are equal in amplitude and $\pi/2$ out of phase. If there is normal mode loss, however, the amplitudes are still equal, but the phase difference decreases in magnitude. Note that this is an example for which merely specifying a power splitting ratio would not capture the coupling matrix, and it illustrates the usefulness of working in the block diagonal form.

Output for general modulation and normal mode loss

To give an example with simplified algebra, we assume that the two couplers have identical normal mode losses as in Eqn (27), and therefore that

$$(\mathbf{C}_n^{(k)})_{11} = c_{nk11} = c_{nk22} = \frac{e^{-j\pi/4}}{2}(1+jp) \quad c_{nk12} = c_{nk21} = \frac{e^{-j\pi/4}}{2}(1-jp). \quad (28)$$

These are the coefficients that are substituted into Eqn (23), in evaluating inner product $\vec{\mathbf{E}}^\dagger \vec{\mathbf{E}}$, and we assume here that they have a common delay which can be ignored. The matrix elements for the waveguides, however, must be attached to the fields they transport, as mentioned above. Specifically, the fields in our fiber 1 are assumed to suffer a time delay of τ with respect to the fields in fiber 2.

Tracking the fields permits an identification by matrix element index set, and when the inner product is taken, those indices indicate the field on which the element operates. Expanding Eqn (23) and treating the terms as operators, we can find operators for the powers in the two fibers as

$$\hat{P}_1 = \frac{(1+p^2)^2}{16} \left[\hat{A} + 2\text{Re} \left(\hat{B} e^{-j4\tan^{-1}p} \right) \right] \quad (29a)$$

$$\hat{P}_2 = \frac{(1+p^2)^2}{16} [\hat{A} + 2\text{Re}\hat{B}] \quad (29b)$$

where

$$\hat{A} = |J_{11}^{(1)}|^2 + |J_{21}^{(1)}|^2 + |J_{11}^{(2)}|^2 + |J_{21}^{(2)}|^2 \quad \text{and} \quad (30a)$$

$$\hat{B} = J_{11}^{(1)\dagger} J_{11}^{(2)} + J_{21}^{(1)\dagger} J_{21}^{(2)} . \quad (30b)$$

The carets in the equations above signify that the scalar matrix elements are to be linked to their respective time-delayed field operands. Once they are so linked, they can be viewed as scalars. After performing this identification of the matrix elements with their time delayed envelopes, we arrive at an expression for the time dependence of the powers exiting the coupler ports:

$$P_1 = \frac{(1+p)^2}{16} [E^*(t-\tau)E(t-\tau) + E(t)^*E(t) + 2\text{Re} [-B e^{+j4\epsilon_p} E^*(t-\tau)E(t)]] \quad (31a)$$

$$P_2 = \frac{(1+p)^2}{16} [E^*(t-\tau)E(t-\tau) + E^*(t)E(t) + 2\text{Re} [B E^*(t-\tau)E(t)]], \quad (31b)$$

where the (nominally small) angle ϵ_p is expanded in p as

$$\epsilon_p = \frac{\pi}{4} - \tan^{-1} p = \tan^{-1} \frac{1-p}{1+p^2} . \quad (32)$$

For an ideal coupler, $\epsilon_p = 0$, and Eqns (31) show that the second terms in the expressions for P_1 and P_2 are complementary, regardless of the relative birefringence in the interferometer's arms. For couplers with normal mode loss, $\epsilon_p \neq 0$, and Eqns(31) show that there is dc phase shift added to the ac signal created by the interfering fields. We illustrate this for the case of a phase modulated input applied to such an interferometer to demodulate it.

Output for phase modulation and normal mode loss

Eqns (31) show that the output powers vary from intensity effects (the $\hat{A}$ operator which gives rise to the E^*E terms) and from interference terms. Here, we investigate the effect of coupler normal mode loss on the demodulated output in a phase modulated link, for which the amplitude is constant. That is, assuming an input field with a phase modulated optical carrier of form

$$E(t) = E_0(t)e^{j(\omega_0 t + \phi_m \sin \Omega t)}, \quad (33)$$

what effect will there be on the demodulator output ports when the links are birefringent and the couplers suffer unequal normal mode losses? The ideal case for a scalar field was investigated by Youngquist *et al.*[15], but we generalize the demodulator to include polarization dependence in the links and couplers.

For phase modulation with an rf tone at Ω , as in Eqn (33), the interference term is given by

$$E^*(t-\tau)E(t) = |E_0|^2 e^{j\phi_m(\sin \Omega t - \sin \Omega(t-\tau))} = |E_0|^2 e^{j2\phi_m \sin \frac{\Omega\tau}{2} \cos(\Omega t - \frac{\Omega\tau}{2})} .$$

Assuming path matching ($\Omega\tau = \pi$), lossless fibers ($A = 2$), and quadrature biasing ($B \sim j$), the powers in the two arms are

$$P_1 = \frac{(1+p^2)^2}{8} |E_0|^2 [1 + \sin(4\epsilon_p + 2\phi_m \sin \Omega t)] \quad (34a)$$

$$P_2 = \frac{(1+p^2)^2}{8} |E_0|^2 [1 - \sin(2\phi_m \sin \Omega t)] , \quad (34b)$$

which, in a balanced detection scheme, would result in a detected power of

$$P_{det} = P_1 - P_2 = \frac{(1 + p^2)^2}{8} [(1 + \cos 4\epsilon_p) \sin(2\phi_m \sin \Omega t) + \sin 4\epsilon_p \cos(2\phi_m \sin \Omega t)] . \quad (35)$$

The net result of this is that the output would appear to be offset from quadrature by approximately $2\epsilon_p$. This effect would be small, but noticeable: for a coupler with a 0.25 dB differential loss, for instance,

$$2\epsilon_p \approx 1 - p \approx \frac{\ln 10}{10} dB_{loss} = 0.058 \text{ rad} ,$$

where ‘ dB_{loss} ’ is the loss one would measure between a single input and the sum of the two outputs of the coupler.

Summary

In this section we have shown how the formalism can include interferometric structures with modulated inputs, extending the usual scalar approach which treats only single tones. We presented a sample calculation showing how the matrix elements can be derived by using linear vector space techniques to exploit known behavior in the normal modes of the system elements, specifically treating normal mode loss for couplers in a phase modulated microwave photonic link.

VI SUMMARY

We have presented a formalism which uses standard linear vector space techniques to analyze interferometric structures with birefringent and lossy elements. By introducing a simple co-ordinate transformation to move between descriptions of the fiber and the coupler elements, we showed that the 4×4 matrices for all of the elements could be reduced to 2×2 diagonal blocks. This simplification permits each of the system components to be treated in the most general way: couplers and waveguides can be treated with the customary two dimensional unitary matrix techniques, and polarization-dependent losses can be introduced with arbitrary pass eigenstates. Because the reduction to the two-dimensional space is so fundamental, and because its formalism covers both couplers and birefringent waveguides, we included an Appendix which derives and summarizes the main results of the theory. The Appendix also provides a brief summary of the geometric representation, which is useful both in visualizing problems and solutions, and in casting the problems into a uniform framework. We illustrated the formalism by analyzing the power distribution at the output ports for a demodulator in a phase modulated microwave link. We found that normal mode loss creates a dc bias phase to the interferometric output and we showed how to estimate that phase.

VII REFERENCES

- [1] Kirkendall, C.K. and Dandridge, A., "Overview of high performance fibre-optic sensing," *J. Phys. D: Appl. Physics*, vol. **37**, pp. R197-R216, 2004.
- [2] V.J. Urick, F. Bucholtz, P.S. Devgan, J.D. McKinney, and K.J. Williams, "Phase modulation with interferometric detection as an alternative to intensity modulation with direct detection for analog-photonic links," *IEEE Trans. Microwave Theory and Techs*, vol. 55(9), pp. 1978-1985, 2007.
- [3] Y.K. Lize, M. Faucher, E. Jarry, P. Oulette, A. Wetter, R. Kashyap, and A.E. Willner, "Low-loss S-, C-, and L-band differential phase shift keying demodulator," CLEO 07, Baltimore, MD, USA, May 2007, Paper CMJJ4.
- [4] A.H. Gnauck and P.J. Winzer, "Optical phase-shift-keyed transmission," *J. Lightwave Tech.*, vol. 23 (1), pp. 115-130, Jan. 2005.
- [5] P.J. Winzer, A.H. Gnauck, C.R. Doerr, M. Magarini, and L.L. Buhl, "Spectrally efficient long-haul optical networking using 112 Gb/s polarization multiplexed 16-QAM," *J. Lightwave Tech.*, vol. 28 (4), pp. 547- 556, Feb. 2010.
- [6] C.-L. Chen and W.K. Burns, "Polarization characteristics of single-mode fiber couplers," *IEEE J. Quantum Elect.*, 1982, **QE-18**, pp.1589-1600.
- [7] Y.K. Lize, J.-C. Richard, P. Samadi, and L.R. Chen, "Polarization dependent formalism of interferometric structures describing DPSK and DQPSK receivers," OSA/CLEO/QELS 2010, paper CThC2.
- [8] Goldstein, D., *Polarized Light*, Marcel Dekker, NY, 2003.
- [9] N.J. Frigo, "A generalized geometrical representation of coupled mode theory," *IEEE J. Quant. Electron.*, vol. QE-22 (11), pp. 2131-2139, 1986.
- [10] J.P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *PNAS*, vol. 97 (9), pp. 4541-4550 (2000).
- [11] N.J. Frigo and F. Bucholtz, "Geometrical representation of optical propagation phase," *J. Lightwave Tech.*, vol. 27 (15), pp. 3283- 3293, Aug. 2009.
- [12] N.J. Frigo, F. Bucholtz, and C.V. McLaughlin, "Polarization in phase modulated optical links: Jones- and generalized Stokes-space analysis," *J. Lightwave Tech.*, vol. 31 (9), pp. 1503-1511 (2013).
- [13] A. Yariv, "Coupled-mode theory for guided-wave optics," *IEEE J. Quantum Elect.*, vol. QE-9, pp.919-933 (1973).
- [14] H.A. Haus, *Waves and Fields in Optoelectronics*, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.
- [15] Youngquist, R.C., Stokes, L.F., and Shaw, H.J., "Effects of normal mode loss in dielectric waveguide directional couplers and interferometers," *IEEE J. Quantum Elect.*, vol. **QE-19**, pp. 1888-1896, Dec. 1983.
- [16] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed., Englewood Cliffs, NJ, Prentice-Hall, 1971.

APPENDIX A

VIII APPENDIX-A: ELEMENTS OF COUPLED MODE THEORY IN TWO- AND THREE- DIMENSIONS

Representation of 2D complex vectors in a real 3D space for telecom applications has been developed in the literature [9, 10, 11, 12], and the main features needed for the development of this paper are summarized here for convenience.

2D equation of motion

Many physical systems can be modeled as a differential equation with a two-dimensional matrix evolution operator as

$$\begin{aligned} \frac{d}{dz} \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \end{bmatrix} &= -j \begin{bmatrix} \bar{k} + \bar{\kappa}_3 & \bar{\kappa}_1 - j\bar{\kappa}_2 \\ \bar{\kappa}_1 + j\bar{\kappa}_2 & \bar{k} - \bar{\kappa}_3 \end{bmatrix} \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \end{bmatrix} \\ &= -j \begin{bmatrix} \bar{\kappa}_3 & \bar{\kappa}_1 - j\bar{\kappa}_2 \\ \bar{\kappa}_1 + j\bar{\kappa}_2 & -\bar{\kappa}_3 \end{bmatrix} \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \end{bmatrix} - j\bar{k} \mathbf{I} \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \end{bmatrix} \end{aligned} \quad (36)$$

where $\mathbf{I}$ is the 2×2 identity matrix and where the average of the diagonal elements, $\bar{k}$, has been identified explicitly. (The variables are barred because we will transform them later into more convenient forms.) In particular, this equation describes both propagation in birefringent media and coupling in 2×2 optical directional couplers. For birefringent propagation, the $\bar{a}_i$ are the field components in two orthogonal polarizations, the $\bar{\kappa}_i$ represent the birefringence, and $\bar{k}$ signifies the overall optical path length. For 2×2 directional couplers, the $\bar{a}_i$ are field amplitudes in the two waveguides, the $\bar{\kappa}_i$ represent the coupling constants between the waveguides and the detuning from phase matching, and $\bar{k}$ represents the overall optical phase. Eqn (36) is “conservative” in the sense that, for real $\bar{k}$ and $\bar{\kappa}_i$, the norm of the 2D vector (assumed to be +1) is preserved. This reduces the vectors to 3 DOF: the components’ relative magnitude, relative phase, and their common overall phase.

Suppressing overall phase

In some situations, the overall phase is unimportant and can be ignored. In such cases, the common phase constant $\bar{k}$ (which may be a function of distance) can be suppressed by transforming [9] to “slowly varying” components as

$$\bar{a}_i(z) \rightarrow e^{-j \int \bar{k}(z) dz} a_i(z) = e^{-j\phi} a_i(z), \quad (37)$$

yielding an equation of motion for those slowly varying components

$$\frac{d}{dz} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -j \begin{bmatrix} \bar{\kappa}_3 & \bar{\kappa}_1 - j\bar{\kappa}_2 \\ \bar{\kappa}_1 + j\bar{\kappa}_2 & -\bar{\kappa}_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}. \quad (38)$$

This equation suppresses the overall phase, makes the operator in brackets a traceless Hermitian matrix, and can be solved for the slowly varying a_i . If desired at the end, the overall phase can be re-introduced by using Eqn (37) to regain the $\bar{a}_i$.

Eqn (38) permits $\vec{\kappa}$ to vary with z , but in what follows, we consider the case for constant $\vec{\kappa}$ to simplify the development.

Pauli matrices

In both the two- and three-dimensional representations of 2D state vectors, the Pauli matrices play a pivotal role, since they connect the 2D and 3D representations, and they form (with the identity matrix) a basis for 2×2 matrices. They are given by

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \sigma_2 = \begin{bmatrix} 0 & -j \\ +j & 0 \end{bmatrix}; \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (39)$$

Any matrix can be expressed in terms of this basis as

$$M = M_0\sigma_0 + M_1\sigma_1 + M_2\sigma_2 + M_3\sigma_3 = M_i\sigma_i = M_0\sigma_0 + \vec{M} \cdot \vec{\sigma}, \quad (40)$$

where the abstract “Pauli vector” is given by

$$\vec{\sigma} = \sigma_1\hat{e}_1 + \sigma_2\hat{e}_2 + \sigma_3\hat{e}_3 = \sigma_i\hat{e}_i, \quad (41)$$

and the expansion coefficients M_i in Eqn (40) are given by

$$M_i = \frac{1}{2} \text{Tr}(\sigma_i M) = \frac{1}{2} \text{Tr}(M \sigma_i). \quad (42)$$

For Hermitian M , the M_i are real, and thus for traceless Hermitian operators, M can be described by a real 3D vector, $\vec{M}$. Therefore, Eqn (38) can be cast as

$$\frac{d}{dz} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -j \begin{bmatrix} \bar{\kappa}_3 & \bar{\kappa}_1 - j\bar{\kappa}_2 \\ \bar{\kappa}_1 + j\bar{\kappa}_2 & -\bar{\kappa}_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -j(\bar{\kappa}_i\sigma_i) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -j\vec{\kappa} \cdot \vec{\sigma} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}. \quad (43)$$

It is useful to express $\vec{\kappa}$ in terms of its magnitude and unit vector:

$$\vec{\kappa} = \bar{\kappa} \hat{\kappa} = \bar{\kappa}(\kappa_1\hat{e}_1 + \kappa_2\hat{e}_2 + \kappa_3\hat{e}_3) = \bar{\kappa} \kappa_i\hat{e}_i = \bar{\kappa}(\sin \theta_0 \cos \phi_0\hat{e}_1 + \sin \theta_0 \sin \phi_0\hat{e}_2 + \cos \theta_0\hat{e}_3), \quad (44)$$

where the last expression puts $\hat{\kappa}$ into polar co-ordinates referenced to the $\hat{e}_i$ basis. This permits the equation of motion, Eqns (38, 43) to be written in either Cartesian or polar co-ordinates for the operator described by $\vec{\kappa}$:

$$\frac{d}{dz} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -j\bar{\kappa}D \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad (45)$$

where D describes unit vector $\hat{\kappa}$

$$D = \begin{bmatrix} \kappa_3 & \kappa_1 - j\kappa_2 \\ \kappa_1 + j\kappa_2 & -\kappa_3 \end{bmatrix} = \begin{bmatrix} \cos \theta_0 & \sin \theta_0 e^{-j\phi_0} \\ \sin \theta_0 e^{+j\phi_0} & -\cos \theta_0 \end{bmatrix} \quad (46)$$

General 2D solution

For fixed $\vec{\kappa}$, the equation of motion for the state, Eqn (45) suggests an eigenvalue problem. Indeed, one can show that the eigenvectors of D can be expressed in terms of the polar co-ordinate angles as

$$|\psi_s\rangle \sim \begin{bmatrix} \cos \frac{\theta_0}{2} e^{-j\phi_0/2} \\ \sin \frac{\theta_0}{2} e^{+j\phi_0/2} \end{bmatrix} \quad |\psi_f\rangle \sim \begin{bmatrix} \sin \frac{\theta_0}{2} e^{-j\phi_0/2} \\ -\cos \frac{\theta_0}{2} e^{+j\phi_0/2} \end{bmatrix}, \quad (47)$$

where the labels signify “slow” and “fast” eigenstates (in accordance with convention for optical propagation in birefringent media) and the eigenvalues $+1$ and -1 , respectively. Dirac notation signifies that we view

the states as abstract vectors, whereas the brackets indicate the states' representations in the chosen basis for our complex 2D space.

For $\vec{\kappa}$ constant, eigenstates (47) can be used to diagonalize Eqn (45). That is, a 2×2 complex matrix S is formed from the two column vectors in Eqn (47), and its adjoint, $S^\dagger$, is transposed complex conjugate. After transforming to the basis formed by the eigenstates, the equation of motion Eqn (45) can be integrated directly to find the unitary transformation matrix relating the state's components at position “ z ” to those at the input as

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_z = \begin{bmatrix} \cos \bar{\kappa}z - j \cos \theta_0 \sin \bar{\kappa}z & -j e^{-j\phi_0} \sin \theta_0 \sin \bar{\kappa}z \\ -j e^{+j\phi_0} \sin \theta_0 \sin \bar{\kappa}z & \cos \bar{\kappa}z + j \cos \theta_0 \sin \bar{\kappa}z \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_0 = \mathbf{C}_{\theta_0, \phi_0}(z) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_0. \quad (48)$$

The matrix operator above is unitary: its inverse is its adjoint. It expresses the three DOF for the operator described by $\vec{\kappa}$ in terms of three angles: θ_0 and ϕ_0 describe the direction of the unit vector $\hat{\kappa}$, and $\bar{\kappa}z$ is related to both the magnitude of $\vec{\kappa}$ and the distance z over which the system evolves.

Recall that this is the solution for constant coefficients. If $\vec{\kappa}$ changes with z , the solution's form is more complicated. However, one can, at least in principle, imagine stepping the solution over segments Δz small enough that the matrix can be considered constant, and constructing a solution out of products of the solution in Eqn (48). Euler's theorem [9],[11] insures that the result is equivalent to a single overall rotation.

Eqn (48) can be expressed more compactly as

$$\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} = \begin{bmatrix} \cos \frac{2\bar{\kappa}z}{2} \sigma_0 & -j \sin \frac{2\bar{\kappa}z}{2} (\hat{\kappa} \cdot \vec{\sigma}) \end{bmatrix} \begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix} = e^{-j(\bar{\kappa}z) \hat{\kappa} \cdot \vec{\sigma}} \begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix}, \quad (49)$$

which clearly shows the effect on the eigenstates: if the initial vector's components $a_i(0)$ are the eigenvectors' components in Eqn (47), then $\hat{\kappa} \cdot \vec{\sigma}$ leaves the vector unchanged, returns ± 1 to the exponent, and the state vector is multiplied by a phase factor $e^{\mp j\bar{\kappa}z}$. In particular, for future reference, we note the special operator given by $\vec{\kappa} = \bar{\kappa}\hat{e}_3$: in this case $\theta_0 = 0$ and Eqn (48) reduces to

$$\begin{bmatrix} a_1(z) \\ a_2(z) \end{bmatrix} = \begin{bmatrix} e^{-j\bar{\kappa}z} & 0 \\ 0 & e^{+j\bar{\kappa}z} \end{bmatrix} \begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix} = \mathbf{C}_{0,0}(z) \begin{bmatrix} a_1(0) \\ a_2(0) \end{bmatrix}. \quad (50)$$

Finally, we make note of the middle expression of Eqn (49), which contains “superfluous” factors of 2 : these enable a direct connection to the geometrical representation, to which we now turn.

Geometric representation

The above has shown that both the matrix operator and its (+1) eigenstate can be characterized by unit 3D vectors. In this section, we review the geometric representations of states and operators. The description of $|\psi_s\rangle$ in Eqn (47) is a particularly useful construction, and we refer to it as the “normal form:”

$$|\psi\rangle \sim \begin{bmatrix} \cos \frac{\theta}{2} e^{-j\phi/2} \\ \sin \frac{\theta}{2} e^{+j\phi/2} \end{bmatrix}. \quad (51)$$

In the form of Eqn (51), the angles θ and ϕ correspond to the polar and azimuthal angles, respectively, of the unit vector representing that state. For $0 \leq \theta \leq \pi$, the relative amplitudes run from unity in state $|1\rangle$ to unity in state $|2\rangle$, and ϕ is the phase difference between the two co-ordinates. It is clear that in Eqn (47), state $|\psi_s\rangle$ is already in normal form while $|\psi_f\rangle$ can be brought into normal form with the change of variables

$$\theta_f = \pi - \theta_s; \quad \phi_f = \phi_s - \text{sgn}(\phi_s).$$

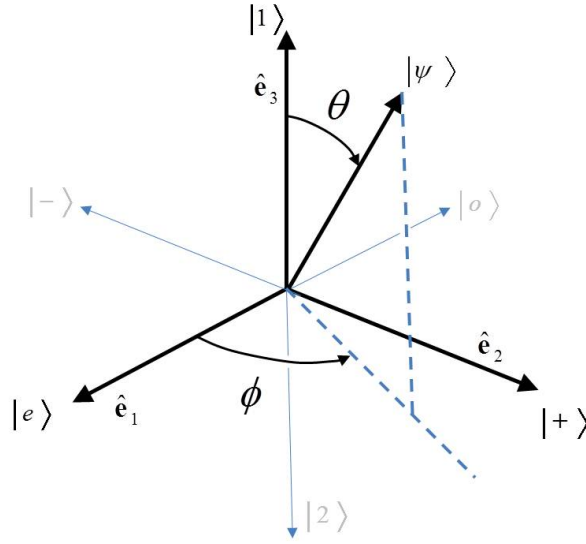


Fig. 8: Representation of states. States in normal form (Eqn (51)) are represented in 3D by plotting components' relative amplitudes via polar angle θ and relative phases via azimuthal angle ϕ . Angles are doubled in this representation, and orthogonal states are on opposite sides of the sphere.

This shows that orthogonal vectors are opposite each other, rather than at right angles, and is an illustration of the fact that in the geometric representation, the angles are twice those of the angles in the 2D representation. An illustration of how states map is shown in Fig. 8.

The Cartesian co-ordinates of the same state $|\psi\rangle$ are given by

$$s_i = \langle s | \sigma_i | s \rangle \implies \vec{s} = \langle s | \sigma_i \mathbf{e}_i | s \rangle = \langle s | \vec{\sigma} | s \rangle, \quad (52)$$

so that, for instance, the component along $\hat{\mathbf{e}}_1$ is

$$s_1 = \langle s | \sigma_1 | s \rangle = \begin{bmatrix} \cos \frac{\theta}{2} e^{+j\phi/2} & \sin \frac{\theta}{2} e^{-j\phi/2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \cos \frac{\theta}{2} e^{-j\phi/2} \\ \sin \frac{\theta}{2} e^{+j\phi/2} \end{bmatrix} = \sin \theta \cos \phi, \quad (53)$$

as expected from the polar co-ordinates in Fig. 8. This is analogous to Eqn (44). By virtue of Eqn (51), and its identification of angles, either polar or Cartesian co-ordinates establish mappings such as

$$\hat{\mathbf{e}}_3 : |1\rangle \sim \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad -\hat{\mathbf{e}}_3 : |2\rangle \sim \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \quad \hat{\mathbf{e}}_1 : |e\rangle \sim \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad \hat{\mathbf{e}}_2 : |+\rangle \sim \frac{1}{\sqrt{2}} \begin{bmatrix} e^{-j(\pi/2)/2} \\ e^{+j(\pi/2)/2} \end{bmatrix}, \quad (54)$$

where we've expressed the two basis states, $|1\rangle$ and $|2\rangle$, their even superposition, $|e\rangle$, and their phase-advanced superposition, $|+\rangle$. While useful, states in the normal form do not show the overall phase necessary to describe interferometric structures.

A state's evolution under an operator, such as in Eqn (48) has an intuitively appealing description in the geometric representation: it is a precession of the state vector about the vector corresponding to the slow eigenstate, $\vec{\kappa}$, at a rate of $2\bar{\kappa}$, shown in Fig. 9. In this figure, an initial state, which corresponds to the 3D vector $\vec{s}_i$, evolves by precessing about the vector $\vec{\kappa}$ at a rate $2\bar{\kappa}$, tracing out a circle on the unit sphere (described in more detail below). We note that: (i) both the operator and the state are represented

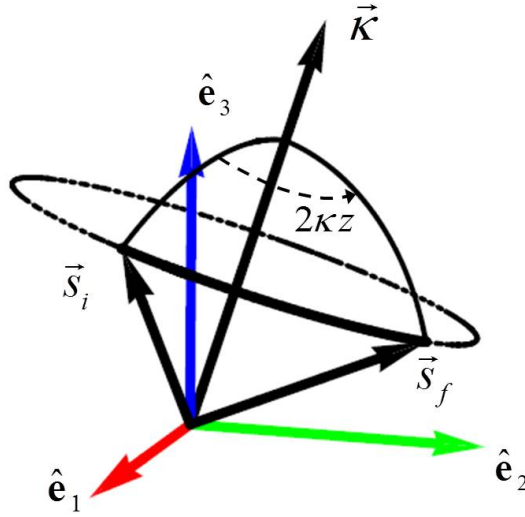


Fig. 9: State evolution. In 3D representation space, the state that corresponds to $\vec{s}_i$ evolves to the state corresponding to $\vec{s}_f$ under the operator corresponding to $\vec{\kappa}$ by rotating about $\vec{\kappa}$ through an angle $2\bar{\phi} = 2\kappa z$, for constant $\vec{\kappa}$.

by vectors in the same 3D space; (ii) the state vector's magnitude doesn't change, since it is being rotated; (iii) the state's rotation rate (and thus the angular displacement) in the 3D representation is twice its rate in the 2D representation; (iv) the arc length of an evolution through angle $2\bar{\kappa}z$ depends on the angular distance between $\vec{\kappa}$ and $\vec{s}_i$: it is a line of latitude when $\vec{\kappa}$ is considered a pole; and (v) in particular, for the two state vectors parallel to, or anti-parallel to, $\vec{\kappa}$, the rotation does not change the vector's position. These two states correspond to the slow (eigenvalue $+1$) eigenstate (along $+\hat{\kappa}$) and the fast (eigenvalue -1) eigenstate (along $-\hat{\kappa}$).

It can be shown[11] that the form of the operator in Eqn (48) is a natural outgrowth of the earlier concepts. That is, consideration of Eqns (50, 51) shows that $C_{0,0}$ operating on $|\psi\rangle$ increases $\phi/2$ by $\bar{\kappa}z$ while keeping θ constant. Thus, operator $C_{0,0}$ corresponds to a precession of the vector $\vec{s}$, representing $|\psi\rangle$ about $\hat{e}_3$. But Eqn (48) is just a similarity transform of $C_{0,0}$ by the S formed from the eigenstates $|\psi_s\rangle$ and $|\psi_f\rangle$, which correspond to $\pm\hat{\kappa}$. Thus, Eqn (48) is the operator which performs (in basis $\pm\hat{\kappa}$) the same operation that $C_{0,0}$ performs (in basis $\pm\hat{e}_3$), namely a precession about $\hat{e}_3$ at rate $2\bar{\kappa}z$. This same result can be derived by direct computation of the transformed differential equation[9].

The summary, to here, is the conventional geometrical description of the coupled mode equations, and has ignored the common phase which was suppressed in going to the slowly varying amplitudes. In the next section, we re-introduce the common phase.

Generalization to vectors with overall phase

Recently we introduced a formalism which permits tracking of the overall phase, the third DOF for a complex vector with unit norm. In addition to the position of the state in 3D, its overall phase is represented by a “fiducial paddle” as shown in Fig. 10. The cardinal states are shown, illustrating that orthogonal states

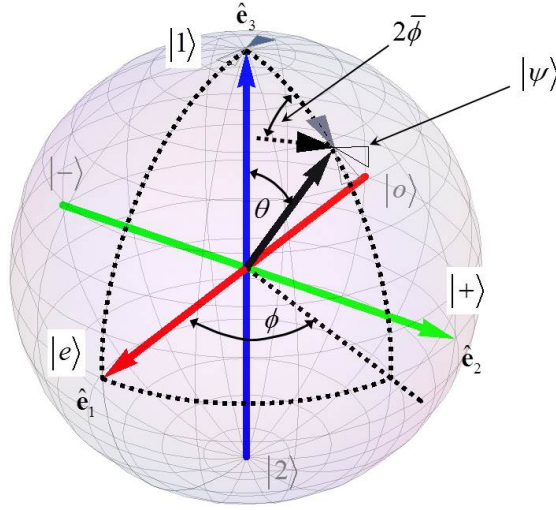


Fig. 10: Representation of phase by fiducial paddles. Arbitrary state $|\psi\rangle$ is represented by the 3D vector $\vec{s}$ and associated fiducial paddle. The relative amplitudes and phases with respect to the 2D bases are shown by $\vec{s}$'s polar and azimuthal angles, θ and ϕ , respectively. The normal form (with zero overall phase) is shown by the shaded fiducial paddle. Overall phase $\bar{\phi}$ is represented by rotating the fiducial paddle through positive (right-hand rule) angle $2\bar{\phi}$.

are on opposite sides of the unit sphere.⁴ An arbitrary state has form

$$|\psi\rangle = e^{-j\bar{\phi}} \begin{bmatrix} \cos \theta/2 e^{-j\phi/2} \\ \sin \theta/2 e^{+j\phi/2} \end{bmatrix} = e^{-j2\bar{\phi}/2} \begin{bmatrix} \cos \theta/2 e^{-j\phi/2} \\ \sin \theta/2 e^{+j\phi/2} \end{bmatrix}, \quad (55)$$

where the right hand expression is written to show the angles as represented in the 3D space. Such a state is shown in the figure to illustrate fiducial paddles, which comprise a colored and a blank blade to show rotation angles. In the figure, the grey paddle (aligned to the line of longitude for ϕ) corresponds to a general state in normal form, Eqn (51). The darker paddle corresponds to the normal form after multiplication by an overall advancing phase factor $e^{-j\bar{\phi}}$: the paddle is rotated in a positive, right-hand rule, sense through angle $2\bar{\phi}$. One envisions that the phase can be counted through multiple rotations of the fiducial paddle. As pointed out earlier [11, 12], the projection of one unit state onto another is effected by moving it (and its fiducial paddle) in parallel transport across a geodesic to that state: the magnitude of the projection is the cosine of half the geodesic angle, and the phase is half the angle between the fiducial paddles. In view of Eqn (51), the arbitrary state Eqn (55), and the half-angle prescription, we can see from Fig. 10 that the projection of general state $|\psi\rangle$ is given by $\langle 1 | \psi \rangle = \cos \frac{\theta}{2} e^{-j(\bar{\phi} + \phi/2)}$.

Generalization to operators with overall phase

Operators which preserve the norm of the vector operands are unitary, so that $U^\dagger U = I$. Since this is a set of 4 equations in the 8 unknowns (4 complex elements of U), there must be 4 DOF for the matrix elements of U . We have seen above that when the overall phase is suppressed, an operator can be expressed

⁴We represent states $|2\rangle$, $|o\rangle$, the odd superposition, and $|-\rangle$, the phase retarded superposition, as “hidden” behind the sphere.

as a 3D vector $\vec{\kappa}$ and an equivalent matrix operator $\mathbf{C}_{\theta,\phi}$. If we *include* the overall phase, Eqn (37) shows that all we need do is multiply the state by $e^{-j\bar{\phi}}$, where $\bar{\phi}$ is the integrated phase up to point z . Then, in light of Eqn (49), unitary operators can be cast in the form

$$\mathbf{U}(\bar{\phi}, \phi, \hat{\kappa}) = e^{-j(\bar{\phi} + (\vec{\kappa} \cdot \vec{\sigma})z)} = e^{-j2\bar{\phi}/2} \left[\cos \frac{\phi}{2} \sigma_0 - j \sin \frac{\phi}{2} (\hat{\kappa} \cdot \vec{\sigma}) \right]. \quad (56)$$

It is clear from this form that U is unitary. In explicit matrix form, this operator is

$$\mathbf{U}(\bar{\phi}, \phi, \hat{\kappa}) = e^{-j2\bar{\phi}/2} \begin{bmatrix} \cos \frac{\phi}{2} - j \sin \frac{\phi}{2} (\kappa_3) & -j \sin \frac{\phi}{2} (\kappa_1 - j\kappa_2) \\ -j \sin \frac{\phi}{2} (\kappa_1 + j\kappa_2) & \cos \frac{\phi}{2} - j \sin \frac{\phi}{2} (-\kappa_3) \end{bmatrix}, \quad (57)$$

where the κ_i are the components of unit vector $\hat{\kappa}$. The effect of this operator is shown in Fig. 11, and can be connected to Fig. 10 in the following way. The spectral theorem [16], in light of Eqn (49), implies that, the eigenstates $\pm \hat{\kappa}$ with eigenvalues ± 1 will pick up a phase difference of $\pm \bar{\kappa}z$. Thus, in a system in which the eigenstates are a basis, the state will rotate about $\hat{\kappa}$ (eigenvalue +1) through angle $\phi = 2\bar{\kappa}z$, picking up relative phase ϕ in *this* basis in a similar way as in Fig. 10. But this is happening with respect to the slow eigenstate $|\hat{\kappa}\rangle$, not $|1\rangle$. At the same time, the state also acquires an overall phase $\bar{\phi}$ from the operator, which further rotates the fiducial paddle by $2\bar{\phi}$ about the $\vec{s}$ axis. As a result, the evolution is depicted as in Fig. 11. Thus, the operator rotates the initial state, vector $\vec{s}_i$ with its fiducial paddle, through angle ϕ to its final state,

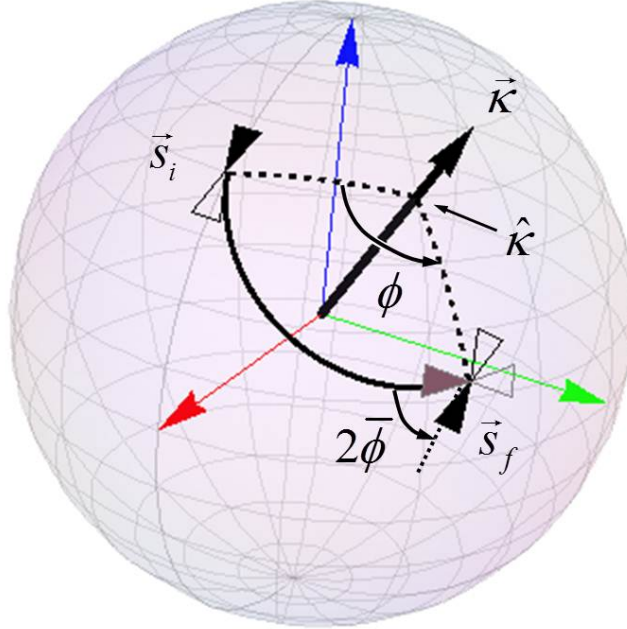


Fig. 11: Geometrical representation of unitary operator. Eigenstate of operator with eigenvalue +1 is $\hat{\kappa}$, with co-ordinates κ_i in Eqn (57). Initial state is represented by $\vec{s}_i$ and a fiducial paddle to represent phase. Operator rotates $\vec{s}_i$ through angle ϕ about $\hat{\kappa}$, while simultaneously rotating the fiducial paddle about $\vec{s}$.

while simultaneously rotating the fiducial paddle about $\vec{s}$'s axis through angle $2\bar{\phi}$.

Application to birefringent propagation and optical coupling

As mentioned above, both propagation in birefringent optical media and evanescent coupling in adjacent waveguides are usually represented with the formalism described in Eqn (36). We briefly review these two applications here.

In optical propagation, $\bar{\phi}$ corresponds to the overall propagation phase, i.e. the average “optical path length” of the medium. The evolution operator represented by the magnitude and unit vector of $\vec{\kappa}$, the medium’s birefringence, which is commonly referred to as $\vec{\beta}$. If $|1\rangle$ is illustratively chosen as the x axis, $\hat{e}_3$, then $\hat{e}_1$ is linearly polarized light at 45° to the x axis, and $\hat{e}_2$ is right-circular polarization. If the “slow” eigenstate were at 45° , and y polarized light were introduced, the state would evolve by precessing about $\hat{e}_1$ in a right-hand sense, from $-\hat{e}_3$ (y polarization) to $\hat{e}_2$ (right circular polarization, to $\hat{e}_3$ (x polarization), to $-\hat{e}_2$ (left circular polarization), and back to $-\hat{e}_3$.

The same sort of evolution occurs for the ideal 3 *dB* coupler. Now, however, the states $|1\rangle$ and $|2\rangle$ correspond to waveguides. For symmetric identical waveguides, the “slow” eigenstate is the even superposition of individual waveguide modes. Thus $\vec{\kappa}$ lies along the $\hat{e}_1$ axis, as above. Light introduced into waveguide 2 would be represented as $-\hat{e}_3$, and the evolution would proceed as described in the paragraph above. In this case, however, the states are interpreted as (i) light in waveguide 2, (ii) light evenly divided, but waveguide 1’s phase advanced by $\pi/2$ with respect to 2’s, (iii) all light in waveguide 1, and (iv) light evenly divided but waveguide 1’s phase retarded by $\pi/2$ with respect to 2’s, and (v) back to all light in waveguide 2.

Conclusion

This Appendix has reviewed state evolution in the basic 2D form and its representation in 3D. We’ve shown that conservative systems are characterized by unitary evolution operators, that the overall phase of an operator can be suppressed and re-introduced at will, that the operator eigenstates are parallel or anti-parallel to the representation of the operator, and that state evolution under a constant operator is given by a precession of the state’s 3D vector about the operator’s vector. The overall phase was shown to be represented by a fiducial paddle which, in addition to being rotated rigidly with the state, is also rotated about the state’s 3D vector by twice the average phase introduced by the operator.