A physical model for geometric phase

by

Luis Alejandro Garza Soto

Copyright © Luis Garza 2024

A Dissertation Submitted to the Faculty of

Engineering

In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

In the Graduate School of

Utsunomiya University

2024

Dedication

I dedicate this thesis to everyone who contributed to my growth as a scientist and as a person. I thank Valeria Tomaselli for her endless love and for being my motivation. I thank my mother and sister for their support throughout the years. I thank my advisor Nathan Hagen for all the hard work he had to go through for accepting me as his student and collaborator, his curiosity and love for science made my research possible.

Table of Contents

Table of Contents—*Continued*

List of Figures

List of Figures—*Continued*

List of Figures—*Continued*

Abstract

Since Pancharatnam's 1956 discovery of optical geometric phase, and Berry's 1984 discovery of geometric phase in quantum systems, researchers analyzing geometric phase have focused almost exclusively on algebraic approaches using the Jones calculus, or on spherical trigonometry approaches using the Poincaré sphere. The abstracted mathematics of the former, and the abstracted geometry of the latter, obscure the physical mechanism that generates geometric phase. This thesis shows that optical geometric phase derives entirely from the superposition of waves and the resulting shift in the location of the wave maximum. This wave-based model provides a way to visualize how geometric phase arises from relationships between waves, and from the transformations induced by optical elements. It also shows how to derive the relationship between the geometric phase of a wave by itself and the phase exhibited by an interferogram, and provide the conditions under which the two match one another. Additionally, this thesis guides the reader through Pancharatnam's original derivation and shows how Pancharatnam's approach connects to the wave description of geometric phase presented. Then, some basic differences between geometric and propagation phase are shown with white light interferometer experiments. The experimental results suggest that some characteristics of geometric phase are described better with help from the physical model presented.

Chapter 1
Introduction

Geometric phase in optics is understood as a shift of the measured phase of light due to changes in it's polarization. Originally studied by Pancharatnam, it remained unpopular for around 30 years. It was until Berry discovered an equivalent phase in the evolution of quantum systems that the concept gained momentum. From that point and forward several researchers dedicated hundreds of papers to observe, calculate, and measure geometric phase. Because of the historical development of the study of geometric phase, it is recognized with different names depending on the context. The phase was first known plainly as Berry's phase. As researchers started observing subtle differences and recognizing Pancharatnam's work, in part thanks to Berry, it became known as Pancharatnam's phase when in the context of polarization optics. Some authors prefer giving credit to both, calling it the Pancharatnam-Berry phase, but this made it unclear if both were the same phase, equivalent, or different phenomena. When the phase is obtained by changes in propagation like those observed by passing a circular state of polarization through a twisted optical fiber, the phase receives the name of spin-redirection phase. The many names it has received and the unclear difference between them perhaps points to a general confusion about the concept and an uncertainty about the mechanism that generates it.

This thesis presents for the first time a physical model for geometric phase, it shows how such model matches the theoretical and experimental developments available until before this research and the reasoning behind it, and it reports on experiments performed to confirm the differences between the geometric phase and the more well known dynamic or propagation phase. In order to familiarize the reader with the terminology that is used throughout the thesis, Chapter 2 introduces basic concepts in polarization optics. This chapter briefly covers the definitions that arise from studying light as a propagating electromagnetic wave. It then presents polarized light and the mathematical formalism used to represent it, how the interaction of matter and light results in changes of polarization, interference of light, and the Poincare sphere. This chapter assumes that the reader has basic knowledge of optics and thus it is used more as a glossary for terminology.

Chapter 3 presents a literature review with a unique perspective. Since this thesis is written after

achieving a more complete understanding of the geometric phase, it would not be fair to explain the literature in retrospective without presenting the results of this thesis first. Thus, this chapter is written attempting to present the knowledge prior to the development of the wave description of geometric phase reported in Chapter 4. The literature review also includes only the papers that were most relevant for the development of the novel ideas presented, despite having studied a much greater amount of papers related to geometric phase.

The wave description of geometric phase, the physical model and main contribution of this thesis, is explained in Chapter 4. First, using basic addition of two waves in 1D, geometric phase is presented as a shift of a composed wave due to the relative displacement and amplitude of its components. The concept is then generalized to the 2D case, where the mathematics are slightly different but the concept is equivalent. A couple of examples are presented to show the reader how the concept applies to light passing through a retarder and how the equations match the results obtained in the literature through other frameworks. Finally, the chapter explains how geometric phase is observed with interference and clarifies how the phase measured and the displacement of the electric field are related.

With a new physical model in hand, it is natural to wonder how it matches the predictions of Pancharatnam and Berry. A precise explanation for the case of Pancharatnam is given in Chapter 5. The difficulty of following Pancharatnam's original paper provided an opportunity to clarify both his explanations and it's relationship to the wave description. The chapter is dedicated to translating his ideas to an easier language for readers.

Chapter 6 emphasizes the differences between geometric phase and propagation phase while clarifying the boundedness problem, whether the geometric phase can take any values or is limited between $-\pi$ and π. The experimental results show why the answer to this question is not as easy as it seems, for both a theoretical and an experimental perspective, and how the answer depends on one's choice of phase convention.

Chapter 2

Basic concepts in polarization

This chapter presents basic concepts in polarization optics that will be useful throughout this thesis. The explanations are derived mainly from two textbooks, "Optics" by Eugene Hecht and "Polarized Light and Optical Systems" by Russell Chipman.[5,6] Repetition in citation will be avoided by acknowledging both books as the references for the whole chapter.

2.1 Electromagnetic waves

Throughout the history of optics the discussion of light being a stream of particles or a propagating wave has been present. Electromagnetic theory and related experiments provided an interpretation of light as a combination of propagating transverse electric and magnetic fields. These fields are generated by a process called electromagnetic radiation where accelerating charged particles cause changes of the electric field distribution in space. The disturbance does not affect all points in space instantaneously, the effect travels in space in the form of a changing electric field.

2.1.1 Propagating waves

The mathematical functions that describe propagating waves of any type are those given by the solutions of the wave equation, including pulses and harmonics. For the electromagnetic case the harmonic solutions can be explained with dipole radiation, where at a distance wavefronts can be considered to be plane and the electric field to change sinusoidally,

$$\mathbf{E} = E_0 \cos\left(\mathbf{k}\cdot\mathbf{r} - \omega t - \phi\right), \tag{2.1}$$

with $\mathbf{k}\cdot\mathbf{r}$ = constant. Note that the wave is expressed in terms of the electric field only, a convention commonly used in polarization optics. The reasoning behind omitting an expression for the magnetic field is that the amplitudes of the fields are known to be proportional and their oscillations orthogonal, allowing to have complete information with the electric field only. In Eq. (2.1) several quantities are introduced, referring to the spatial and temporal dependencies of the wave.

2.1.2 Frequency and wavelength

At a specific time, the wave is defined in different positions in space. The magnitude of the electric field at these positions is clearly also different. The harmonic oscillation of the electric field allows us to define quantities that describe its' cycles. The distance from beginning to end of a cycle, often described by peaks of the sinusoidal function, is defined as the wavelength (λ). The amount of oscillations that can be counted per unit of distance, a sort of spatial frequency, is known as the wavenumber (k). Considering a count of the oscillations per unit distance in a specific direction, we define the k-vector, ($\mathbf{k}$). Note that the argument of the cosine in Eq. (2.1) should have no units. It follows that the product of position $\mathbf{r}$ (m in international units) and k-vector $\mathbf{k}$ (1/m in international units) does not either, and that it results in a description of the electric field's magnitude at different positions. On the other hand if we focus in a specific point in space, we would observe an increasing and decreasing electric field as the wave passes by. The time it takes to go through one full cycle of these values is the period (T). The amount of oscillations that can be counted per unit of time is known as the (temporal or angular) frequency (ω). In this case, the product of time t (s in international units) and frequency ω (1/s in international units) has no units and in the argument of Eq. (2.1) it describes the temporal evolution of the electric field magnitude.

2.1.3 Phase

The space and time dependent arguments inside the cosine function representing the electric field can be accompanied in general by a phase ϕ to represent an initial offset. Also, when a wave is displaced due to other effects like modifying its path it is common to consider that in this term.

2.1.4 Transmission and reflection

When a disturbance is propagating through a medium, for example a mechanical wave on a string, it may reach a boundary where there is a change in the medium. Depending in the change in conditions part of the wave will go through the next medium, the transmitted wave, and part will return through the original medium, the reflected wave. This is also true for electromagnetic waves and some characteristics will be explained after the introduction of other basic concepts.

2.1.5 Superposition

For most waves we observe that when two waves travel and meet at some position, the total effect is a sum of the individual effects. Mathematically, we say that superposition is equivalent to linear addition of the functions describing the waves. A common example of superposition is standing waves: when two transverse waves of the same frequency oscillating along the same axis travel in opposite directions, the addition results in a stationary pattern. However, in polarization optics we are mostly interested in the superposition of two or more electric field waves that propagate in approximately the same direction. Superposition is important both for the definitions of polarized light and for the phenomena known as interference, both covered in the following sections.

2.2 Polarized light

As light propagates its electric field vector traces a path and we say light is polarized when there is a preferred oscillation and traced path. From superposition we know that any electric field can be expressed as an addition of two or more parts. Then, for a beam of light propagating in any direction, the electric field vector at any moment in time can be expressed as the addition of three electric fields oscillating in the x, y, and z directions.

2.2.1 Jones vectors

When we consider that a beam of light is of a single wavelength (monochromatic) and propagates in the z direction, we can describe it using a Jones vector. These vectors have two complex elements or components, one for each direction x and y, and each component has a corresponding amplitude (A_x, A_y) and phase (ϕ_x, ϕ_y). The general Jones vector,

$$\mathbf{E} = \begin{bmatrix} A_x e^{-i\phi_x} \\ A_y e^{-i\phi_y} \end{bmatrix}, \tag{2.2}$$

represents a state of polarization (SOP) of light, where the negative argument of the exponential allows for a positive displacement of the components when ϕ_x or ϕ_y are positive. When the Jones vector is scaled to a magnitude of 1 ($A_x^2 + A_y^2 = 1$) we say it has been normalized, a case that is often used for examples. The Jones vector is also used to trace the polarization ellipse, an iconic representation of polarized light, by multiplying by the temporal phase $e^{-i\omega t}$ and taking the real part.

2.2.2 Linearly polarized light

When the components of the electric field have a phase difference $(\phi_x - \phi_y)$ of 0 or 180 degrees the resulting path traced by the tip of the arrow of the electric field vector is a line. When the amplitude of the y component (A_y) is 0 the state of polarization is horizontal. The phase ϕ_x turns then into a global phase that strictly speaking is multiplying both components but retains no physical meaning where the amplitude is 0. The electric field in this case is $\mathbf{E} = A_x e^{-i\phi_x}\mathbf{H}$, where

$$\mathbf{H} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.3}$$

represents the normalized horizontal SOP with no global phase. The opposite condition, $A_x = 0$, yields $\mathbf{E} = A_y e^{-i\phi_y}\mathbf{V}$, where

$$\mathbf{V} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{2.4}$$

Both $\mathbf{H}$ and $\mathbf{V}$ are examples of linear states of polarization that in general oscillate along a line oriented at an angle ψ with respect to the horizontal,

$$\psi = \tan^{-1}\left(\frac{A_y}{A_x}\right), \tag{2.5}$$

given by the relationship between the sides of the triangle formed by the electric field vector and its' components. When both amplitudes are the same, $A_x = A_y = A$, and both phases are the same, $\phi_x = \phi_y = \phi$, the linear SOP oscillates along the 45° direction. Any linear SOP can be expressed in terms of an orientation angle, amplitude, and global phase,

$$\mathbf{E} = Ae^{-i\phi}\begin{bmatrix} \cos(\psi) \\ \sin(\psi) \end{bmatrix}. \tag{2.6}$$

From this point forward this thesis refers to any of the normalized linear states of polarization as "ψ degrees linear SOP".

2.2.3 Circularly polarized light

When both components of the electric field have the same amplitude and a phase of $\pi/2$ between them, the total electric field keeps its amplitude at all times and rotates tracing a circle. If we consider normalized amplitudes and all of the phase to be in the y component, $\phi_x = 0$, $\phi_y = \pi/2$, the corresponding Jones vector is

$$\mathbf{E} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}. \tag{2.7}$$

This is one of two SOPs known as circular states of polarization, the other one corresponding to an opposite rotation of the electric field vector,

$$\mathbf{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}, \tag{2.8}$$

where the components have shifted in order, $\phi_y = -\pi/2$, but kept the same phase between them. Depending on a choice of convention for the direction of rotation, it's helicity, we name these two SOPs right circular polarization (RCP) and left circular polarization (LCP), discussion presented briefly after the introduction of elliptically polarized light.

2.2.4 Elliptically polarized light

The more general combination of the parameters A_x, A_y, ϕ_x, and ϕ_y, results in tracing some ellipse. The polarization ellipse shown in Fig. 2.1 is characterized by the length of the major axis a, the length of the minor axis b, and the ellipticity ϵ,

$$\epsilon = \frac{b}{a}, \tag{2.9}$$

which refers to how flat or round the ellipse is. The ellipse also has an orientation angle ψ that depends on the amplitudes, A_x, A_y, and the relative phase between the components, $\phi = \phi_y - \phi_x$. The arrow tip in the ellipse represents the initial condition of the electric field in time and it gets displaced along the ellipse if the beam of light has a global phase, a phase shared by both components. It is most common to define a state of polarization from the addition of horizontal and vertical components, but this is arbitrary. Any pair of orthogonal oscillations (referred to as a basis in this thesis) can be used as the components of the beam and when different basis are used the global phase is different.

2.2.5 Phase sign conventions

The argument of the cosine function in Eq. (2.1) can be both positive or negative and represent the same wave because the cosine function is symmetrical,

$$\cos(\theta) = \cos(-\theta). \tag{2.10}$$

This makes it possible to adopt one of two conventions. Either phase increases in space and decreases with time, $\frac{2\pi}{\lambda}\mathbf{k}\cdot\mathbf{r} - \omega t - \phi$, or the phase decreases in space and increases with time, $\omega t - \frac{2\pi}{\lambda}\mathbf{k}\cdot\mathbf{r} + \phi$.

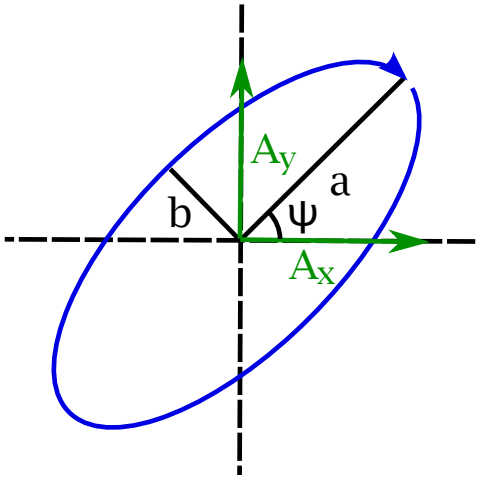

FIGURE 2.1. The polarization ellipse showing the major axis (a), minor axis (b), and orientation angle (Ψ) of the ellipse.

Throughout this document we will consider the former. With such definition Eq. (2.7) represents right circularly polarized light and Eq. (2.8) represents left circularly polarized light.

2.3 Polarization elements

The analysis of a beam of light's state of polarization is done using materials that have optical properties like absorption and refraction that depend on polarization. A material is dichroic when it's absorption coefficient depends on the polarization state. Such behavior was first observed on a gemstone and later replicated in polarizers, optical elements that are designed to obtain polarized light as an output. A material is birefringent when it's refractive index depends on the polarization state. Some crystals like calcite and quartz naturally exhibit birefringence and they can be cut and polished to create retarders, also known as waveplates.

2.3.1 Diattenuation and retardance

The previously mentioned polarization dependent absorption is measured in terms of transmitted intensity. The diattenuation D,

$$D = \frac{T_{max} - T_{min}}{T_{max} + T_{min}} \tag{2.11}$$

is a metric for the strength of a polarizer, and takes values between 0 and 1. An element that transmits all polarization states equally has a diattenuation of 0. An ideal polarizer, which com-

pletely blocks one state of polarization and completely transmits another state of polarizarion, has a diattenuation of 1.

In the case of birefringent crystals, two states of polarization remain unchanged upon propagation. These are known as the eigenpolarizations of the material. An incident polarized beam of light can be expressed as the superposition of two beams with oscillations corresponding to the eigenpolarizations. These two beams, which we also call components, travel different optical path lengths (OPL) and the optical path difference (OPD) of these components is defined as the retardance δ. Retarders are linear, circular, or elliptical depending on their eigenpolarizations, and their axes are named "fast axis" for the SOP corresponding to the smaller OPL and "slow axis" for the SOP corresponding to the larger OPL.

2.3.2 Jones matrices

The mathematical representation of elements that modify the SOP of light are 2x2 matrices known as Jones matrices. Together with Jones vectors they comprise Jones calculus, a system for polarization calculation. The Jones matrix relates an arbitrary incident polarization state $\mathbf{E_0}$ to the corresponding exiting polarization state $\mathbf{E'}$ by matrix-vector multiplication,

$$\mathbf{E'} = \mathbf{J} \cdot \mathbf{E_0} = \begin{bmatrix} j_{xx} & j_{xy} \\ j_{yx} & j_{yy} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} j_{xx}E_x + j_{xy}E_y \\ j_{yx}E_x + j_{yy}E_y \end{bmatrix} = \begin{bmatrix} E'_x \\ E'_y \end{bmatrix}. \tag{2.12}$$

The eigenvectors and eigenvalues of a Jones matrix determine the properties of the optical element. The eigenvectors correspond to the eigenpolarizations. The eigenvalues have the information of the diattenuation in their amplitudes, and of the retardance in the difference of their phases. A Jones matrix is homogeneous if it's eigenpolarizations are orthogonal. Thus, when referring to homogeneous optical elements (polarization elements) in this thesis we use this definition and not the perhaps more common one referring to spatially homogeneous optical properties.

2.3.3 Polarizer and retarder Jones matrices

The Jones matrix for a horizontal linear polarizer,

$$\mathbf{LP}(0^\circ) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{2.13}$$

transmits all of the x-component of the incident light and bocks the y-component. The Jones matrix for a linear polarizer (LP) with transmission axis oriented at an angle θ,

$$\mathbf{LP}(\theta) = \begin{bmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{bmatrix}, \tag{2.14}$$

can be calculated using the Jones matrix for the horizontal linear polarizer multiplied by the corresponding rotation matrices. For a linear retarder oriented horizontal, a different phase is obtained by each component,

$$\mathbf{LR}(\delta, 0) = \begin{bmatrix} e^{-i\phi_1} & 0 \\ 0 & e^{-i\phi_2} \end{bmatrix}, \tag{2.15}$$

where the difference between this phases is the retardance of the linear retarder, $\delta = |\phi_1 - \phi_2|$. Since for most polarization purposes only retardance is relevant, it may be split in any way between the components. There are three standard conventions for such splitting: slow axis unchanged, fast axis unchanged, and symmetric. The symmetric convention, used throughout this thesis, considers half of the retardance to be a positive phase for one component and the other half to be a negative phase for the other component. The other conventions consider phase to be obtained completely by one component or the other. Using rotation matrices on Eq. (2.15), the Jones matrix for a linear retarder of retardance δ oriented at an angle θ,

$$\mathbf{LR}(\delta, \theta) = \begin{bmatrix} e^{-i\delta/2}\cos^2(\theta) + e^{i\delta/2}\sin^2(\theta) & -i\sin(\frac{\delta}{2})\sin(2\theta) \\ -i\sin(\frac{\delta}{2})\sin(2\theta) & e^{i\delta/2}\cos^2(\theta) + e^{-i\delta/2}\sin^2(\theta) \end{bmatrix}, \tag{2.16}$$

is obtained. The two most commonly used linear retarders are those that have a retardance of $\pi/2$ and π which correspond to one quarter of a wave and half a wave respectively. These retarders receive the names of quarter wave plate (QWP) and half wave plate (HWP).

2.4 Interference

Interference refers to the measurable effects that occur when two or more light waves are combined. A series of conditions determine if two beams of light are coherent or incoherent, referring to being able to interfere or not respectively. When two beams are coherent they can interfere constructively or destructively, creating interference patterns. The study of this patterns can reveal important information about the optical properties of a material and can even be used to measure diverse physical quantities.

2.4.1 Interferometers

Interferometers are tools or setups used to obtain the mentioned interference patterns. The main idea behind interferometers is dividing a beam of light trough some mean and combining the parts to obtain interference. The division of the beam could be done either by passing parts of a wavefront through different slits or windows (Young's double slit interferometer), or by reflecting and transmitting some portion using a beamsplitter (division of amplitude).

Michelson

The Michelson interferometer uses a beamsplitter to separate light in two paths. Each beam is then reflected on a mirror and travels in the opposite direction. Both beams combine after passing through the beamsplitter for a second time. Light exits the beamsplitter for the second time in two directions, back to the source and opposite to one of the mirrors. A screen or camera is used in the latter and some interference is observed.

Mach-Zehnder

The Mach-Zehnder interferometer also splits light in two paths using a beamsplitter. Each beam reflects on a mirror at some angle (different configurations may be used for specific purposes) and then both meet at a second beamsplitter. The beams are combined in both outputs and the Mach-Zehnder interferometer allows for the observation at both. The interference pattern depends on a small angle between the propagation vectors of the beams. An advantage of the Mach-Zehnder interferometer is that light propagates through the arms of the interferometer only once and in only one direction. This allows for simpler analysis of the state of polarization and it's evolution upon interacting with the optical elements.

Sagnac

The Sagnac interferometer is special in the sense that it is a single path interferometer. Light is transmitted and reflected by a beamsplitter creating two beams. Mirrors are set up such that both beams travel the same path in opposite directions. This is useful to avoid phase changes of the interferogram due to OPD caused by vibrations or movement of elements.

2.5 Stokes Parameters and the Poincaré sphere

The Jones vectors presented in section 2.2 do not represent unpolarized light or partially polarized light. To characterize the polarization properties of light beams that are unpolarized or partially polarized, the Stokes parameters are used. These parameters are often organized in "Stokes vectors" which are useful for calculation but are not true vectors because they do not transform or rotate as vectors. The Stokes parameters are defined in terms of six polarized flux measurements performed using linear polarizers and linear retarders.

2.5.1 Polarized flux measurements

The intensity measurement of a beam of light after passing through a linear polarizer oriented at either 0°, 45°, 90° or 135° is a horizontal (P_H), 45 degree (P_{45}), vertical (P_V), or 135 degree (P_{135}) polarized flux respectively. Additionally, RCP polarized flux (P_R) and LCP polarized flux (P_L) are measured with circular polarizers which can be obtained by stacking a LP and a HWP with a $\pm 45°$ angle between them. The Stokes parameter S_0 is a measurement of the total intensity of a beam of light, $S_0 = P_H + P_V$. The Stokes parameter S_1 compares how much of the beam of light is horizontal and how much is vertical, $S_1 = P_H - P_V$. When light is polarized completely horizontal $S_1 = 1$ and when it is polarized completely vertical $S_1 = -1$. The Stokes parameter S_2 compares how much of the beam of light is 45° polarized and how much is 135° polarized, $S_2 = P_{45} - P_{135}$. The Stokes parameter S_3 compares how much of the beam of light is RCP and how much is LCP, $S_3 = P_R - P_L$.

2.5.2 The Poincaré sphere

The Poincaré sphere is a geometric representation of SOPs and their Stokes parameters that simplifies the analysis of some polarization problems. The normalized Stokes parameters are used as the three axis defining the space in which the sphere exists and each point on the sphere represents a different SOP. It was through the study of the Poincaré sphere that Pancharatnam found one of the most influential anticipations of the geometric phase, concept introduced through the literature review in chapter 3.

Chapter 3

Literature review

This chapter presents a literature review that does not attempt to explain the geometric phase from a pedagogic perspective. It rather shows the developments perceived as most relevant for the development of this thesis in chronological order. While some review papers explain the topic through the author's perspective using the literature as support, this chapter briefly explains some of the relevant papers available with a perspective prior to the development of my research. As shown below, the geometric phase has been found in several areas of physics and even if a variety will be briefly covered, the group of papers chosen for this literature review heavily favor geometric phase in optics. Reviews with a different scope and points of view, like those that focus on geometric phase as a quantum effect, are available in the literature.[7]

3.1 Generalized theory of interference and its applications. Part I. Coherent pencils

In 1956, S. Pancharatnam was working with absorbing biaxial crystals and observed interference when passing light through them.[8] His search for a description of the phenomena brought him to defining two beams of light with different SOPs to be in phase when their interference is maximum, concept later recognized as the Pancharatnam connection.[9] With great mathematical prowess Pancharatnam found a geometric relationship between SOPs on Poincaré sphere and the phase of the interference pattern. His research was basically ignored for many years, perhaps because of the difficulty of making sense of such relationship or his writing style.

3.1.1 Solid angle approach

When separating a beam of light in two beams with different SOPs the amplitudes should be split such that the intensity of interference equation holds. The SOP of the original beam and the SOP of the two resulting beams can be represented as points on Poincaré sphere. Those points can then be connected with geodesics, a reasonable choice considering Pancharatnam's definitions. The phase between the two resulting beams is then given by the solid angle subtended between the geodesics.

If we consider the three SOPs in Pancharatnam's explanation to be an evolution of the SOP of a beam of light, we can obtain the geometric phase of the evolution using the solid angle approach. For example, some system modifies the SOP of a beam of light from horizontal to diagonal to RCP and back to horizontal, as shown in Fig. 3.1. The solid angle subtended within the geodesics is then 1/8 of the whole sphere. The area of a unitary sphere is 4π and it follows that the solid angle and the geometric phase in this case is $\pi/2$. When we consider an evolution through less convenient SOPs like three different elliptical SOPs, the calculation of the solid angle turns more complicated and associating some physical meaning for geometric phase is not possible.

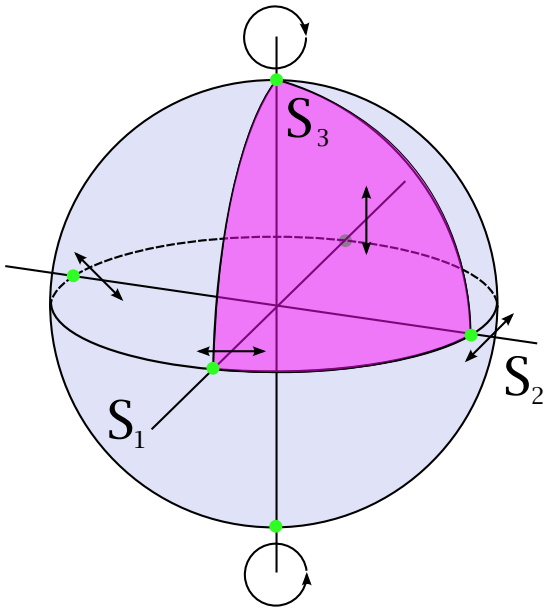

FIGURE 3.1. Solid angle between the geodesics joining the points on Poincaré sphere corresponding to horizontal, diagonal, and right circular polarization.

3.2 Quantal phase factors accompanying adiabatic changes

In 1983 M. V. Berry theorized about the adiabatic evolution of quantum states.[10] He explains that if the environment and the corresponding Hamiltonian $\hat{H}$ are slowly altered, the system will be in an eigenstate of the instantaneous $\hat{H}$. From this it follows that when the system returns to its original state there is an acquired phase factor. Berry shows that such phase factor contains a circuit dependent component dependent. In the last section of the paper the physical effects of magnetic vector potentials[11] are related to the geometrical phase factor.

3.3 Observation of Berry's topological phase by use of an optical fiber

In 1986 Tomita and Chiao reported the first experimental verification of Berry's phase.[12] The authors performed an experiment proposed earlier[13] where a linear SOP is passed through a coiled optical fiber and that results in an extra optical rotation. They also mention the effect is not caused by torsional stress but rather arises from the geometry of the path taken by light. The fiber was twisted into a helix in two ways, first with a uniform or constant pitch helix and then with deformations, proving that as long as the solid angle of the path in momentum space stayed constant the geometric phase would be the same.

3.4 Observation of topological phase by use of a laser interferometer

In 1988 Bhandari and Samuel reported on an experimental observation of the Pancharatnam phase (different terms were used for geometric phase in optics or in other contexts).[1] This was the first experimental observation of geometric phase with light that focused on evolution of states of polarization rather than on an optical rotation. The mechanism by which the SOP evolves is by passing through wavelates and polarizers. The different SOPs in which light is before and after going through the elements are connected through geodesics like the ones used by Pancharatnam to find his relationship for geometric phase. Their results showed how the measured phase matched the one predicted by theory, specifically by the solid angle.

3.5 A geometric-phase interferometer

In 1992 Hariharan and Roy presented an interferometer operating purely on geometric phase.[3] Their setup uses a He-Ne laser as light source and the beam of light passes through a linear polarizer oriented at 45°. The beam then passes through a polarizing beamsplitter giving two orthogonally polarized (horizontal and vertical) beams as an output. The beams then reflect on two mirrors aligned so that a triangular common path is formed, corresponding to a triangular Sagnac interferometer. Between these two mirrors a set of waveplates composed of two QWPs with the same orientation and a HWP between them that is free to rotate, known as a QHQ, is used to increase geometric phase. Since both beams travel in opposite directions, the orientation of the equivalent retarder that describes the QHQ would get flipped in the chosen perspective (convention).

However, for Hariharan's QHQ the eigenvectors resembled the 0° and the 90° orientations and such flipping is avoided. The authors provide a graphical description of the geometric phase obtained by drawing the geodesics between the SOPs the beams have when propagating through the system. The solid areas shown are equal and the direction of rotation is opposite. The amount of geometric phase obtained by passing through the elements should then be equal in magnitude and opposite in sign for both beams. Their results show how the interferogram obtained matches the theoretical predictions.

3.6 The Pancharatnam-Berry phase for non-cyclic polarization changes

In 2010 T. van Dijk et al showed a setup that allowed for observation of geometric phase obtained when the SOP evolution does not finish at the original state (non-cyclic polarization changes).[14] Their approach heavily relies on using Jones calculus to calculate the geometric phase and allows the reader to follow the procedure. Experiments show how the geometric phase measurements are compared to the calculated values for open and closed paths. A relevant contribution of this paper is that it shows that geometric phase can increase linearly and non-linearly depending on the path followed and in it being open or closed.

3.7 Pancharatnam-Berry phases of optical systems

In 2011 Gutierrez-Vega presented simple closed-form expressions for evaluating the Pancharatnam-Berry phase introduced by an optical system.[15] The author combines mathematical concepts of spin matrices and Jones calculus to represent general optical elements. Analyzing the eigenvectors and eigenvalues of the system Gutierrez-Vega is able to predict the geometric phase introduced by a single element mathematically, a concept that is still cause of confusion for come researchers. His expressions additionally consider parallel transport when multiple elements are stacked. Geometrical analysis is also presented through figures that show a combination of the geometric phase obtained by passing through each element and parallel transport. The explanation is mathematically accurate but difficult to follow and the physical meaning of geometric phase is not clarified.

3.7.1 Jones calculus approach

The same example used for the solid angle approach in section 3.1.1 (horizontal-diagonal-RCP-horizontal) can be used for the Jones calculus approach. One way to generate the evolution of SOPs discussed is to pass a horizontally polarized beam of light through a HWP at 22.5°, then through a QWP at 67.5°, and finally through a QWP at −45°. This set of transformations is equivalent to the multiplication of the Jones matrices of each element and the Jones vector of the input beam of light,

$$\mathbf{E}_{out} = \mathbf{LR}(\pi/2, 45°) \cdot \mathbf{LR}(\pi/2, 90°) \cdot \mathbf{LR}(\pi, 22.5°) \cdot \mathbf{E}_{in}.$$

The calculation is not complicated but it is lengthy and can be performed with some computational tool. The result of the calculation is the original SOP multiplied by a global phase. Since the Jones matrix for every element neglects thickness (and so it neglects propagation phase) the total phase obtained is geometric phase. The use of a computational tool facilitates calculation but if we want an analytical solution the problem is very demanding and interpretation is difficult. Once again, this approach does not allow us to associate a physical meaning to geometric phase.

3.8 Geometric phase morphology of Jones matrices

In 2017 Lopez-Mago et al demonstrated a technique based on the Pancharatnam-Berry phase (one more name geometric phase has received) that can determine the geometric phase of an optical system.[16] Their setup uses a linear polarized He-Ne laser that passes through a HWP to obtain 45° linear polarization. The beam then passes through a beam displacer resulting in two orthogonally, horizontal and vertical, polarized beams propagating parallel to each other. The beams pass through a combination of waveplates, first a QWP and then a HWP, that can be rotated with motorized mounts. The combination of the orientation angles of the QWP and the HWP can cause the input SOP to change to any other SOP, thus this set is called a state generator by the authors. The two beams with general orthogonal SOPs are the light source of a Mach-Zehnder interferometer with linear retarders in one arm and no elements in the other. Their idea is that by passing two orthogonal SOPs through the same system, the geometric phase obtained by each beam is equal in magnitude and opposite in sign. At the camera a set of horizontal fringes is observed per pair of beams, both displaced vertically in opposite directions due to the optical system. The subtraction

of their vertical positions depends only on geometric phase since the dynamic phase is equivalent for both sets of fringes. The authors provide geometric phase maps for different optical systems including a basic example of an homogeneous optical system (QHQ) and an inhomogeneous optical system (QWP-LP-QWP).

3.9 Geometric phase polarimetry

In 2020 Garza-Soto et al report two interferometric techniques that use geometric phase to find the eigenvectors and eigenvalues of an optical system.[17] The paper shows how the expressions found by Gutierrez-Vega and later on demonstrated by Lopez-Mago can be used to find the eigenvectors and eigenvalues of a system.[15,16] The author derived a relationship between visibility of interference fringes and geometric phase that can be used to substitute a phase measurement that is susceptible to vibrations with a visibility measurement that is not affected by dynamic phase.

Chapter 4

Wave description of geometric phase

As we have come to learn more about geometric phase over the past 38 years,[8,10] we have acquired the ability to calculate it for an increasing array of circumstances. Yet, even now we lack a physical model for visualizing geometric phase. For polarization optics, the Poincaré sphere has been a widely used tool to visualize the various calculations and to describe how the geometric phase relates to transformations of polarization states, but its use opens up even more questions.[18] For example, when using the Poincaré sphere to calculate the geometric phase, the typical definition states that when a state of polarization (SOP) undergoes a series of transformations, the various transformations induce a phase delay γ to the wave, equal to half the solid angle subtended on the surface of the sphere: $\gamma = \Omega/2$. However, when drawing the curves used to delineate the spherical area, researchers often ignore the fact that the actual physical path taken by the polarization state is not that of the curves they draw. Thus, while the calculation proceeds correctly, the underlying physics that drives the calculation methods remains unclear.

The following presents a list of some of the curious rules under which the Poincaré sphere areas must be calculated in order to obtain correct estimates of the geometric phase:

1. Pancharatnam's original work never considered a cycle of states, but rather considered only two input states A and B, and the state generated by their sum, $C = A + B$.[8] The states A and B, together with C' (the antipode to C), form a triangle of points on the Poincaré sphere from which one can calculate the subtended solid angle, Ω.[19]

2. The subtended angle formed by the actual physical path of the polarization state is generally not the correct solid angle Ω needed to get the correct geometric phase. Rather, for the correct Ω one must use the shortest geodesic arc connecting the pair of SOPs before and after each homogeneous polarization element.[18,20] In addition, if the physical path of the polarization state *is* a geodesic, such as when using a half-wave plate (HWP), then one must use the geodesic that coincides with the physical path.

3. One of the consequences of separating the physical path of the polarization state from the

geodesic path of calculation is that it becomes difficult to determine the value of the geometric phase inside optical elements such as a linear retarder. Rather, analyses are generally limited to calculating γ before entering or after leaving an element, but not the continuous changes that occur while propagating through it.

4. The areas on the Poincaré sphere are actually signed areas — negative if clockwise, positive if anticlockwise, and that if a path crosses itself then one can have positive and negative areas partially cancelling one another.[21–23]

5. While the spherical angle is usually defined using a closed cycle of polarization states,[7,24] a cycle of states can never in practice be exactly closed, it is necessary to have a procedure for calculating the geometric phase that allows for a set of states that are only approximately closed, or not closed at all.[14] This procedure is as follows: the phase of non-closed loop of states $A \rightarrow B \rightarrow C$ is equal to the Pancharatnam–Berry phase related to the closed loop $A \rightarrow B \rightarrow C \rightarrow A$, plus the phase of the Pancharatnam connection[25] related to the projection of $A \rightarrow C$.[22]

6. Under some conditions, the geometric phase can undergo a π shift singularity.[14,20,26,27]

While the above rules are useful guides to obtaining the correct result, the abstraction of the spherical geometry leaves little insight into the physical origin guiding why these rules must exist, and can even hinder understanding. Definitions found in the existing literature, such as "Geometric phase is a consequence of parallel transport in a curved topology"[28] or similar abstract treatments[29] are not exactly wrong, but they can be considered misleading, in that much simpler processes are actually at work. In the discussion below, we show that simply analyzing the locations of wave maxima, without the abstraction of the Poincaré sphere or of matrix algebra, generates the geometric phase properties while also allowing one to visualize the underlying physics. Moreover, in contrast to existing methods, our wave-model of the geometric phase provides a clear means of defining the geometric phase at any point in its propagation through an optical system — not only at points before and after traversing a homogeneous optical element, but inside the element as well.

Section 4.1 starts with the simplest possible case — the superposition of waves in 1D — and introduces a reference plane from which we can define geometric phase. When two waves are superposed, the location of the wave amplitude peak of their sum depends on both the relative

phases of the input waves and on their relative amplitudes. This change in position from the input reference plane to output state peak location is the geometric phase.

Section 4.2 generalizes this superposition analysis to 2D, and shows how we can define the phase of an arbitrarily polarized wave with respect to the two orthogonally polarized waves that compose it. This section also introduces a visualization aid for locating the polarized wave peaks.

In Sec. 4.3, we provide an example of propagating polarized waves through a linear retarder. Although previous authors have claimed that single linear retarders cannot introduce geometric phase,[21] more recent work has argued that indeed they can,[15] and our work as well shows that such retarders do in general cause a global phase shift.

Finally, we note that the above definitions of geometric phase involve propagating waves, and so these are theoretical constructs that are not directly measurable. Section 4.5 introduces a model for interfering a sample wave with a reference wave, generating a stationary interferogram that is measurable. The resulting interferogram includes the geometric phase γ previously obtained for propagating waves. However, one finds that if the two arms of the interferometer deliver waves of different polarization states, then the interferogram phase will differ from the wave geometric phase. This is similar to the requirement sometimes stated in the historical literature that the geometric phase is only defined for "closed loop" transformations of the polarization state. However, we show that this condition is in fact too restrictive. Rather than requiring that the two beams share the same polarization state, we find that the two beams need only share the same ellipticity.

4.1 Wave composition in 1D

It is a well-known property of sinusoidal waves that the sum of any two waves is also a sinusoid, with an amplitude and phase that depends on the states of the input waves. For two co-propagating electromagnetic waves E_1 and E_2 of arbitrary amplitude and phase $E_1 = A_1 \cos(kz - \omega t - \phi_1)$ and $E_2 = A_2 \cos(kz - \omega t - \phi_2)$, their sum is given by the Harmonic Addition theorem as[5,30]

$$E_1 + E_2 = A_3 \cos(kz - \omega t - \gamma)\,, \tag{4.1}$$

where $k = 2\pi/\lambda$ is the wavenumber, ω the angular frequency, and where

$$A_3^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\phi_2 - \phi_1)\,, \tag{4.2}$$

$$\tan\gamma = \frac{A_1 \sin\phi_1 + A_2 \sin\phi_2}{A_1 \cos\phi_1 + A_2 \cos\phi_2}\,, \tag{4.3}$$

give the resultant amplitude A_3 and phase γ. This can also be generalized to the case of adding N arbitrary waves, as shown in Appendix 4.7.1. Note that by using cosines to represent our two waves, we have implicitly chosen the wave peaks to indicate the phase origin ($\phi = 0$) position. A phase advance is given by $\phi > 0$, and a phase delay by $\phi < 0$.

In order to make the wavelength-dependence of Eqs 4.2 & 4.3 explicit — something that will be important when we do this derivation for 2D waves in Sec. 4.2 — we can rewrite each phase as $\phi = 2\pi\xi/\lambda$ for a given physical shift distance ξ.

If we choose to locate our reference plane halfway between the E_1 and E_2 peaks, the phase difference $\delta = \phi_2 - \phi_1$ between the two waves is split in half on each side of the reference plane, such that $\phi_1 = \delta/2$ and $\phi_2 = -\delta/2$. In this case, the phase γ of the sum wave simplifies to

$$\tan\gamma = \tan(\delta/2)\frac{A_1 - A_2}{A_1 + A_2}\,. \tag{4.4}$$

As indicated by our choice of variable, γ is the geometric phase for this superposition.[31] If the amplitudes of the two input waves are equal ($A_1 = A_2 = A$), using the product identity for cosines with (4.2) obtains

$$E_3 = 2A\cos(\delta/2)\cos(kz - \omega t)\,. \tag{4.5}$$

Thus, when the two input wave amplitudes are equal, the sum wave amplitude becomes $A_3 = 2A\cos(\delta/2)$ and the sum wave phase is exactly at the midpoint between the phases of the input waves ($\gamma = 0$). The value of γ indicates the location of the sum wave peak relative to the reference plane. If the two input wave amplitudes are not equal ($A_1 \neq A_2$), then Eq. (4.4) indicates that the phase of the sum wave shifts towards the phase of the input wave with greater amplitude. This situation is illustrated in Fig. 4.1, where we see that the phase of the green sum wave moves closer to the blue wave phase due to the latter's higher amplitude.

In (4.4), one might be concerned about the case when $A_1 = -A_2$, in which case the denominator in the equation becomes zero. However, our choice of the phase delay between the two component waves allows the negative sign to be replaced with an extra phase delay of $\delta \to \delta + \pi$, while maintaining positive amplitudes. This avoids possible division by zero.

If the phase separation δ between the two input waves is greater than π, then an ambiguity arises. Since we cannot generally measure the absolute phase of a wave, but can only measure phase differences between waves, when $\delta > \pi$ the nearest E_2 peak (red wave) with respect to the E_1 (blue wave) peak will no longer be to the left of the origin, but rather to the right. In our

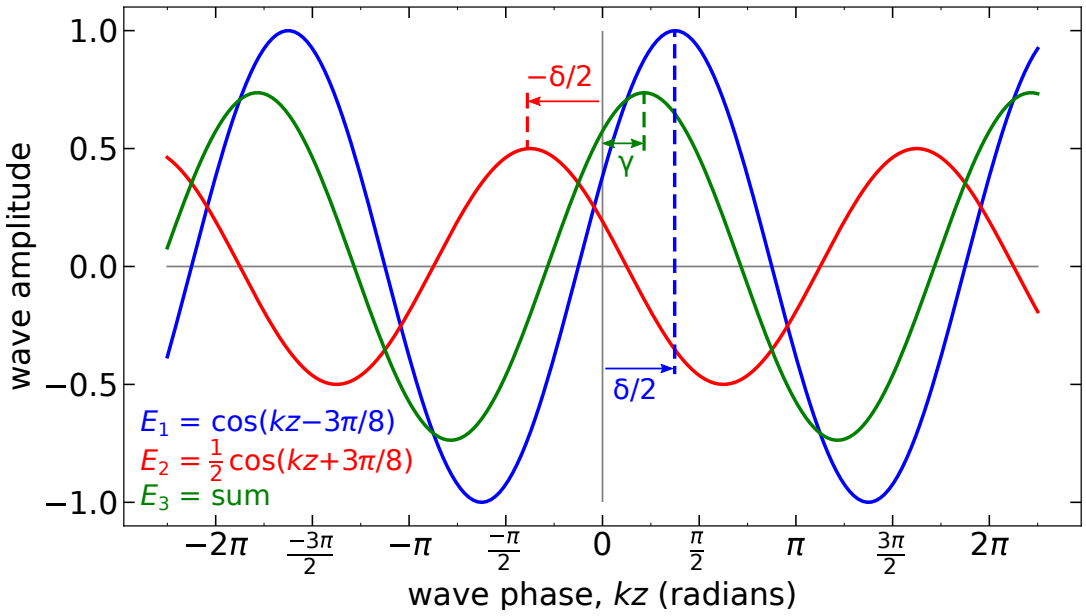

FIGURE 4.1. Superposition of two waves with amplitudes $A_1 = 1$ (blue) and $A_2 = 0.5$ (red) and a phase difference $\delta = 3\pi/4$ between them. The peak location γ of the resultant wave (green) shifts towards the wave with greater amplitude (blue). The phase of each wave ($\phi_i = \pm 3\pi/8$) is defined with respect to a reference position ($z = 0$ here) given by the midpoint between the two component wave peaks. The phase $\gamma = 38.8°$ of the sum wave's peak is given by (4.4).

convention, we always choose the nearest pair of peaks, so that δ for E_3 (green wave) can never exceed π. As we will see below, this change from using the left peak to the right peak corresponds to behavior discussed in the existing literature on the geometric phase, in which the geometric phase measured in interferograms is wrapped within a range of $-\pi$ and $+\pi$.

4.2 Wave composition in 2D

Unlike the 1D case, we cannot simply add orthogonally-vibrating waves and look for the resulting peak. Instead, we form the polarization ellipse traced by the electric vector produced by adding the two orthogonal waves. The point at which the electric vector aligns with the major axis of the ellipse is designated to be the location of the 2D wave peak — its phase. Generalizing our approach to plane waves in 2D involves the same basic procedure as in 1D, but we will see that we must work with the squares of the electric fields rather than the fields themselves.

To start, we use the midpoint between the two orthogonal component waves' peaks to define a reference plane, and define the geometric phase as the phase difference of the elliptical wave's sum peak from the reference plane. While this is easy to do when working with linear polarization

states, waves in 2D are elliptically polarized in general — their instantaneous electric field vectors have an orientation that varies with time. Some polarization components are more conveniently represented using a basis of two elliptical states, but if we want to define a "phase" for representing an elliptically polarized wave, we need a convention. Pancharatnam proposed the natural choice of defining $\phi = 0$ at the position where the electromagnetic vector is at its maximum positive displacement from the axis. This is also known as *Pancharatnam's connection*, and can also be stated as defining the relative phase ϕ between two interfering waves such that $\phi = 0$ (the two waves are "in phase") when their interference is maximally constructive.[32]

For a general elliptical state of polarization propagating along the z-axis, the real-valued electric fields of the E_x and E_y components are given by

$$E_x = A_x \cos(kz - \omega t - \phi_x)\,, \tag{4.6}$$

$$E_y = A_y \cos(kz - \omega t - \phi_y)\,. \tag{4.7}$$

At any given position z and time t the vector magnitude of the sum of these two components is

$$|E| = \sqrt{A_x^2 \cos^2(kz - \omega t - \phi_x) + A_y^2 \cos^2(kz - \omega t - \phi_y)}\,. \tag{4.8}$$

In order to locate the z-position at which the electric field magnitude is maximum, we take the derivative with respect to z (at constant t) and search for the location where the derivative is zero. After some algebraic work, the derivative becomes

$$\begin{aligned}\frac{d|E(z)|}{dz} = &-kA_x^2 \sin\left[2(kz - \omega t - \phi_x)\right] \\ &- kA_y^2 \sin\left[2(kz - \omega t - \phi_y)\right]\,.\end{aligned} \tag{4.9}$$

We can see that this has a similar form to (4.1), and so it comes as little surprise that after some algebra we obtain a similar expression for γ:

$$\tan(2\gamma) = \frac{A_x^2 \sin(2\phi_x) + A_y^2 \sin(2\phi_y)}{A_x^2 \cos(2\phi_x) + A_y^2 \cos(2\phi_y)}\,. \tag{4.10}$$

We can see that the expression for γ for 2D waves closely resembles that for 1D waves (4.3), but that the amplitudes are now squared, and the ϕ's are all doubled. The phase doubling is expected, since the squaring of the electric fields causes the wavelength to halve.

Thus, we can say that a geometric phase of, say, $\gamma = 90°$ in the case of 2D waves corresponds to exactly the same z-location as a 45° phase in the 1D case. These two cases indicate the same

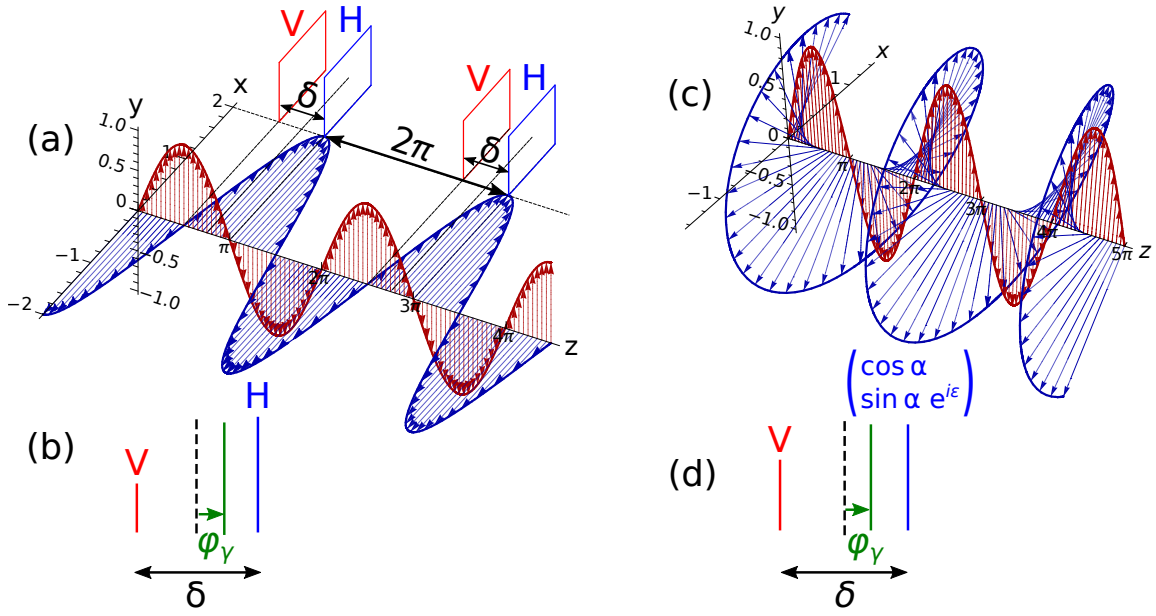

FIGURE 4.2. (a) Two waves oscillating in the x and y axes, with $A_x = 2$, $A_y = 1$, and $\phi_x = \pi$, $\phi_y = \pi/2$. The H and V "wavefronts" are drawn at the wave peaks. (b) The wavefronts from (a). (c) A vertically polarized and an elliptically polarized wave, with amplitudes $A_1 = 1$, $A_2 = 1.5$ and phases $\phi_1 = \pi/2$, $\phi_2 = 5\pi/8$. (d) The wavefronts from (c).

distance, but in terms of different waves, and thus the factor of two difference in phase value. Keeping track of different wavelengths in expressions becomes confusing, and so one should always keep in mind that the "oscillation wave" (what we call the sinusoidal modulation in the electric field vector length) has half the wavelength, and that we always use ϕ to represent the phase expressed in terms of electric field wavelength λ and not the oscillation wavelength $\lambda/2$.

As a result of this 2D composition of waves, we can say that any 2D polarized wave possesses a geometric phase value. And, as we show in App. 4.7.2, the (4.10) agrees with existing expressions in the geometric phase literature obtained via Jones calculus.[16]

As with wave superposition in 1D, if there is no fixed reference plane, then it is convenient to choose the reference plane to be such that ϕ_x and ϕ_y are symmetrically displaced with the reference halfway between them: $2\phi_x = -2\phi_y = \delta$. In such a case, (4.10) simplifies to

$$\tan\gamma = \tan(\delta/2)\,\frac{A_x^2 - A_y^2}{A_x^2 + A_y^2}\,. \tag{4.11}$$

Comparing the 2D Pancharatnam phase (4.11) with the 1D Pancharatnam phase (4.4), we can see that they have almost the same form, but with the 1D expression's amplitudes A_1, A_2 replaced with the 2D expression's A_x^2, A_y^2.

Figure 4.2(a) illustrates two co-propagating linearly-polarized waves, with electric field amplitudes $A_x = 2$ and $A_y = 1$. At the location where each wave reaches its maximum, we draw a plane to indicate the wave position — the "wavefront". For phase difference $\delta \neq 0$, the superposition of these two waves will be an elliptically polarized wave (not shown in the figure) that has its maximum displacement at γ with respect to the midpoint reference plane (the dashed line). The location of the sum-wave's peak is indicated by γ.

Figure 4.2(b) shows a simplified drawing where the wavefronts are represented by vertical lines. The length of each line indicates the amplitude of the wave, and the horizontal distance between the two lines represents the phase difference δ between the two components. The reference plane is indicated by a vertical dashed line, given by the midpoint between the peaks of the two component waves.

Figure 4.2(c) shows the case of superposing a vertically polarized and an elliptically polarized wave. Once again, the wave positions are defined with respect to the field vector maxima. (For the elliptical wave, this is maximum is equal to half the length of the polarization ellipse's major axis.) The corresponding wavefront representation is shown in Fig. 4.2(d).

Figure 4.3 shows three examples how this wavefront representation can be used to help visualize wave composition and the resulting geometric phase γ. Figure 4.3(a) shows a horizontal polarized wave and a vertical polarized wave that are in phase and have the same amplitude. Their sum produces diagonally polarized light (azimuth angle 45°) with respect to the x axis. From (4.10), we find that the position of the wavefront of this sum wave is given by $\gamma = 0$ since the relative phase between the components is $\delta = 0$. This means that the diagonally-polarized sum wave is in phase with both of the input component waves.

Figure 4.3(b) shows a horizontally polarized wave and a vertical polarized wave with equal amplitudes, and a phase difference $\delta = \pi/2$. The addition of the two waves produces right-circularly polarized light. This time (4.10) gives $\gamma = 0$ because the amplitudes of the two component waves are equal. Once again, the sum wave is in phase with the inputs.

Figure 4.3(c) shows the general case for summing two waves using an x-y basis. The two input waves have different amplitudes $A_{\mathrm{H}} \neq A_{\mathrm{V}}$, and a phase difference δ. The sum wave will have a phase γ lying between the two input wavefront positions.

While Eqs 4.3 & 4.10 give the expression for the geometric phase, it does not yet represent a physically measurable quantity, because an interferogram is needed to detect a global phase shift.

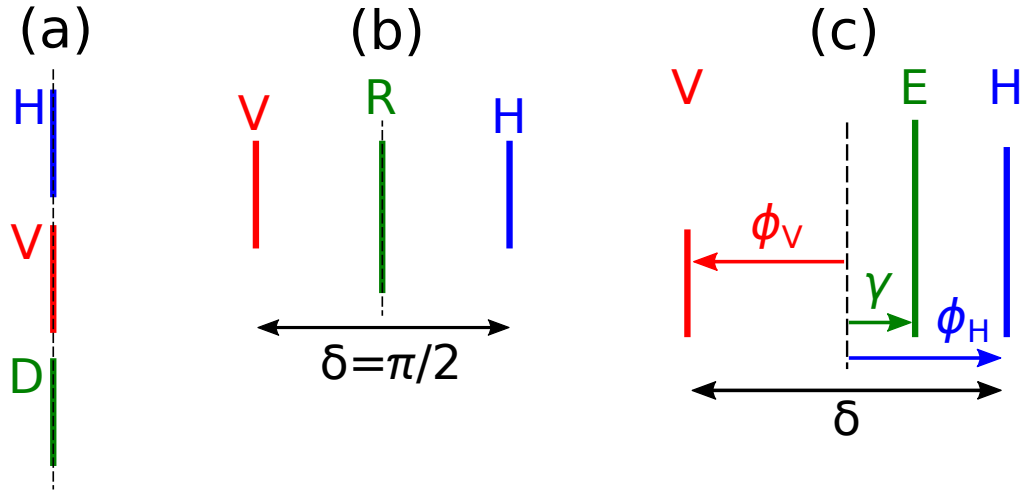

FIGURE 4.3. Using the wavefront representation to aid wave composition calculations. (a) Input horizontally-polarized (H) and vertically-polarized (V) waves sum to create a diagonally-polarized (D) wave oriented at 45°, in phase with the two input waves. (b) The same input waves as in (a) but now with a phase difference $\delta = \pi/2$ between them. The resultant sum wave is elliptically polarized in general, but right-circularly polarized (R) if $A_H = A_V$. Since the two input waves are in phase with one another, the sum wave is in phase with them. (c) Input waves of arbitrary amplitudes A_H and A_V and phase difference δ. The resultant elliptical sum wave (E) has position γ relative to the midpoint between the wavefront positions of the two input waves.

Section 4.5 discusses how this expression can be generalized for the addition of 2D waves to produce a measurable quantity. However, we can note that the expression given in (4.11) contains much of the behavior of geometric phase that excites curiosity. For example, if the waves were to pass through a retarder R, and we want to compare the input wave phase to the output wave phase, we cannot simply subtract their phases because the retardance induced by R will be present inside the tangent function of (4.11). This induces a nonlinear behavior in the phase, such that the path by which one reaches the output has an effect on the calculation.

4.3 The effect of polarization components on geometric phase

In order to demonstrate how our wave-based approach to geometric phase can deal with optical elements, we use our wavefront visualization technique (Fig. 4.3) and analyze the effect of a linear retarder on the wavefront phase.

Figure 4.4(a) shows each step for calculating the phase for a horizontally-polarized wave propagating through a linear retarder whose fast axis is oriented at 45°:

Step 1 The input wave is horizontally polarized, so the V-wave component is zero.

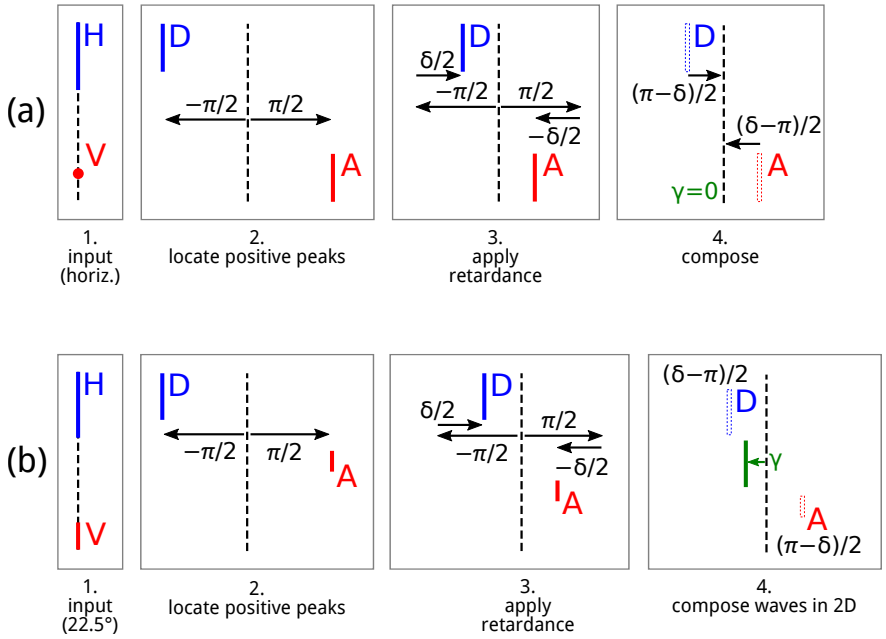

FIGURE 4.4. Diagram for visualizing the propagation of a polarized wave through a linear retarder (retardance δ) with its fast axis oriented at 45°: (a) a horizontally-polarized (H) wave input, (b) a 22.5°linearly-polarized wave input.

Step 2 Transform the horizontally-polarized wave into the eigenbasis of the retarder: the diagonal and antidiagonal (D,A) linear polarization states. Since a projection from a linear polarization state onto any orthogonal linear polarization basis produces only 0 phase shift, the D and A waves have to be symmetrically shifted with respect to the reference plane. While it is also possible to use a $-\pi$ phase shift for the A-wave, this creates a distance of more than π between positive wave peaks. Our convention is to shift all wavefronts such that phase differences between the positive wave peaks being added never exceeds π.

Step 3 Apply the retardance to the two waves, by advancing the D wave and delaying the A wave symmetrically. (The propagation phase due to the retarder is incorporated into the dynamic phase and, as we will see in Sec. 4.5, does not play a role in the final result.)

Step 4 Compose the D and A waves to determine the phase of the resulting elliptical state. Since the peak of the resulting state coincides with the location of the original reference plane, we find $\gamma = 0$.

In this example the D and A waves superposed after passing through the retarder are of equal amplitude. As a result, the phase of their sum is located at the midpoint between the two input

peaks, and therefore $\gamma = 0$. In fact, for a 45° orientation linear retarder, this situation of zero geometric phase will occur for any input polarization state that is located on the red circle drawn in Fig. 4.5.[16]

In order to demonstrate the method by more conventional means, we follow the same five steps above with numerical calculations. Step 1 begins with a polarization state $\vec{E}_{\rm H} = (1\ 0)^{\rm T}$. In order to apply the waveplate retardance, we need to decompose this into the (D,A) basis. Doing so gives two waves, $\vec{E}_1$ and $\vec{E}_2$ given by

$$\vec{E}_1 = \frac{1}{\sqrt{2}}\vec{E}_{\rm D}\,, \quad \vec{E}_2 = -\frac{1}{\sqrt{2}}\vec{E}_{\rm A}\,, \tag{4.12}$$

where $\vec{E}_{\rm D}$ and $\vec{E}_{\rm A}$ are normalized Jones vectors for the D and A polarization states. Once we apply the retardance (Step 3), these two waves become

$$\vec{E}_1 = \frac{1}{\sqrt{2}}e^{-i\delta/2}\vec{E}_{\rm D}\,, \quad \vec{E}_2 = -\frac{1}{\sqrt{2}}e^{i\delta/2}\vec{E}_{\rm A}\,. \tag{4.13}$$

Here we have ignored the dynamic phase — an assumption that is justified later in (4.25).

In Step 4, we compose the D- and A-wave into a single 2D polarized wave using (4.10), with

$$\bar{A}^2 = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{-1}{\sqrt{2}}\right)^2 = 1 \tag{4.14}$$

and

$$\bar{\gamma} = \arctan\left[\frac{\sin(\delta/2) + \sin(-\delta/2)}{\cos(\delta/2) + \cos(-\delta/2)}\right] = 0\,, \tag{4.15}$$

where the overbars indicate that the amplitude and phase refer to the 2D wave rather than to either of its 1D component waves. (Note that the equal amplitudes of D- and A-waves have allowed the amplitudes to factor out from the expression.)

From 4.15, we see that the geometric phase $\bar{\gamma}$ produced by the waveplate in this case is zero — a result that holds true for any input polarization state that lies on the red circle drawn in Fig. 4.5. For an input SOP that does not lie on the red circle, the geometric phase will not be zero. This result agrees with the previous findings of Ref. 16.

The step-by-step process outlined above produces the same geometric phase as is obtained from the standard Jones vectors.[15] One drawback of the Jones vector approach is that so far it has only been applied to calculate γ before and after homogeneous optical elements, and not within them. In this approach, for the situation shown in Fig. 4.4(a), the Jones vectors of the input and output

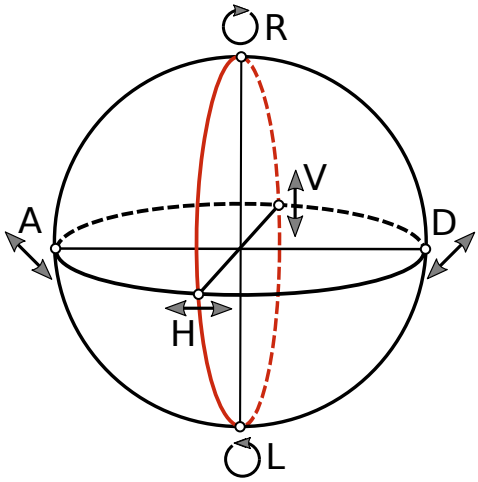

FIGURE 4.5. The great circle drawn in red on the Poincaré sphere indicates the states of polarization for which $\gamma = 0$ via 4.10 (i.e. when composing waves in the D-A basis.) Dashed curves indicate locations that are on the opposite face of the sphere. The axis labels (H, V, D, A, R, L) indicate the horizontal, vertical, diagonal, antidiagonal, right-circular, and left-circular polarization states.

waves are $\vec{E}_a = (1 \;\; 0)^{\mathrm{T}}$ and

$$\vec{E}_b = (1/\sqrt{2})\,(1 \;\; 1)^{\mathrm{T}} \tag{4.16}$$

respectively. With these two, the geometric phase obtained from the retarder is given as

$$\gamma = \arg\left\{\vec{E}_a^{\dagger}\vec{E}_b\right\} = 0\,. \tag{4.17}$$

From this we can see that all three methods — the graphical approach of Fig. 4.4, the mathematical approach of (4.10), and the Jones-vector approach — agree on the result.

4.4 The effect of polarization components on geometric phase, II

Figure 4.4(b) shows a second example, this time of a 22.5°linear polarization state propagating through the same linear retarder as in Fig. 4.4(a). This time, because the amplitude splitting is no longer symmetric, we will have a geometric phase even when the element is a quarter-wave plate.

Step 1 In the input wave, the H-wave component is smaller than in Fig. 4.4(a), and the V-wave component is not zero.

Step 2 We project the 22.5°polarization state onto the (D,A) eigenbasis for the retarder.

Step 3 Apply the retardance: $+\delta/2$ to the D-wave, $-\delta/2$ to the A-wave.

Step 4 Compose the D- and A-waves into a single elliptically-polarized wave with peak location given by γ.

For the 22.5°input case, we see that the final wavefront position shifts towards the component wave with the larger amplitude, and this location is in general shifted by γ from the initial reference plane. This is a "geometric" phase because it does not depend on propagation distance z or time t, but rather depends only on the relationships between the input SOP, the eigenbasis of the retarder, and the retardance.

As with Sec. 4.3, we can make the graphical approach above quantitative as follows. When we decompose the initial 22.5°polarized wave into the (D,A) basis, we obtain

$$\vec{E}_1 = \cos(\pi/8)\vec{E}_{\rm D}\,, \quad \vec{E}_2 = -\sin(\pi/8)\vec{E}_{\rm A}\,. \tag{4.18}$$

Once we apply the retardance, these become

$$\vec{E}_1 = \cos(\tfrac{\pi}{8})e^{-i(\delta-\pi)/2}\vec{E}_{\rm D}\,, \quad \vec{E}_2 = \sin(\tfrac{\pi}{8})e^{-i(\pi-\delta)/2}\vec{E}_{\rm A}\,. \tag{4.19}$$

The elliptical wave created by composing these two has amplitude $\bar{A}$ and phase shift $\bar{\gamma}$ given by

$$\bar{A}^2 = \cos^2(\pi/8) + [-\sin(\pi/8)]^2 = 1 \tag{4.20}$$

and

$$\begin{aligned}\tan\bar{\gamma} &= \frac{\cos^2(\frac{\pi}{8})\sin([\pi-\delta]/2) - \sin^2(\frac{\pi}{8})\sin([\pi-\delta]/2)}{\cos^2(\frac{\pi}{8})\cos([\pi-\delta]/2) + \sin^2(\frac{\pi}{8})\cos([\pi-\delta]/2)} \\ &= -\frac{\tan[(\delta-\pi)/2]}{\sqrt{2}}\,.\end{aligned} \tag{4.21}$$

Therefore, for a quarter waveplate retarder ($\delta = \pi/2$), we find that the geometric phase is $\bar{\gamma} = -0.615$ radians.

We also also follow the Jones vector method, for which the input and output waves $\vec{E}_{\rm a}$ and $\vec{E}_{\rm b}$ are

$$\vec{E}_a = \begin{pmatrix}\cos(\pi/8)\\ -\sin(\pi/8)\end{pmatrix}\,, \quad \vec{E}_b = \begin{pmatrix}\cos(\pi/8)e^{-i(\delta-\pi)/2}\\ \sin(\pi/8)e^{+i(\delta-\pi)/2}\end{pmatrix}\,. \tag{4.22}$$

For $\delta = \pi/2$, these give

$$\gamma = \arg\left\{\vec{E}_a^{\dagger}\vec{E}_b\right\} = -0.615\,\text{radians}. \tag{4.23}$$

In this example, the input polarization state does not split in equal amplitudes between the two eigenstates of the retarder. As a result, we find that the geometric phase is no longer zero, but that the peak of the elliptical wave emerging from the retarder is shifted by -0.615 radians with respect to the input reference plane.

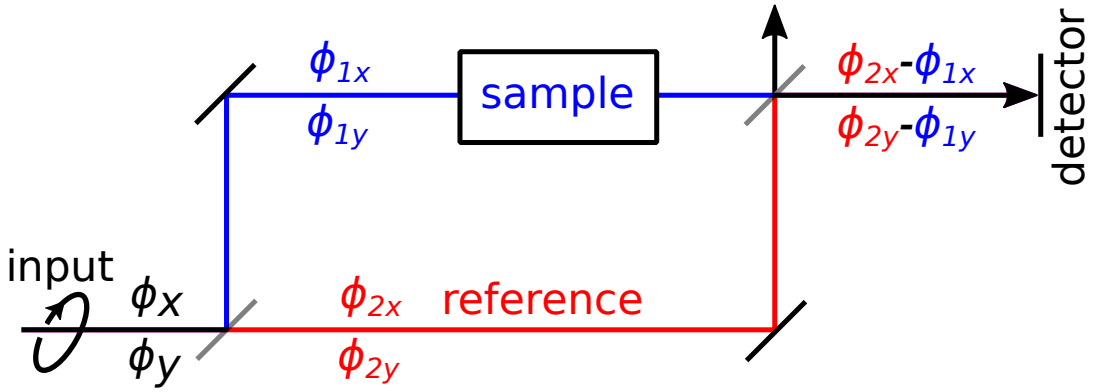

FIGURE 4.6. A Mach-Zender interferometer, showing the wave phases used in Eqs 4.24–4.25.

4.5 Interferogram detection of geometric phase

Measurement of optical geometric phase requires an interferometer, in order to detect the geometric phase as a shift with respect to a reference phase. This involves combining the output from the sample arm of the interferometer — the one for which we are calculating the geometric phase — with a wave transmitted through the reference arm of the interferometer. Figure 4.6 shows the optical layout for a Mach-Zender interferometer — the model we use below for considering interferometric measurement. In a typical setup, the polarization state of the reference arm of the interferometer is designed to match the SOP of the light output by the sample arm.

We write wave $\vec{E}_1$ as the electromagnetic wave from the sample arm, and $\vec{E}_2$ as the wave from the reference arm, using the Jones vectors

$$\vec{E}_1 = e^{i\psi_1}\begin{pmatrix} A_{1x}e^{-i\phi_{1x}} \\ A_{1y}e^{+i\phi_{1y}} \end{pmatrix}, \quad \vec{E}_2 = e^{i\psi_2}\begin{pmatrix} A_{2x}e^{-i\phi_{2x}} \\ A_{2y}e^{+i\phi_{2y}} \end{pmatrix}, \tag{4.24}$$

where $\psi = kz - \omega t$ is the propagation phase (also called the "dynamic phase") obtained by the wave after passing through the interferometer, and ϕ encodes the phase delay between the x and y components of the wave.[17] At this point, it is likely not at all clear that the sample arm encodes a geometric phase, but App. 4.7.3 shows how the expression for $\vec{E}_1$ in (4.24) can be expressed in a form to make the geometric phase $e^{i\gamma}$ explicit. The phase difference $\Delta = \psi_2 - \psi_1$ is the difference in optical path length between the two arms of the interferometer — a degree of freedom that exists in the interferometer setup.

When the sample and reference waves combine to produce the interferogram, the resulting

intensity distribution is expressed by

$$\begin{aligned} I &= \left(\vec{E}_1 + \vec{E}_2\right)^\dagger \left(\vec{E}_1 + \vec{E}_2\right) \\ &= A_{1x}^2 + A_{2x}^2 + 2A_{1x}A_{2x}\cos[(\psi_1 - \psi_2) - (\phi_{1x} - \phi_{2x})] \\ &\quad + A_{1y}^2 + A_{2y}^2 + 2A_{1y}A_{2y}\cos[(\psi_1 - \psi_2) - (\phi_{1y} - \phi_{2y})] \\ &= C + 2A_{1x}A_{2x}\cos(\Delta - \delta_x) + 2A_{1y}A_{2y}\cos(\Delta - \delta_y)\,, \end{aligned} \tag{4.25}$$

where $C = A_{1x}^2 + A_{2x}^2 + A_{1y}^2 + A_{2y}^2$, $\delta_x = \phi_{1x} - \phi_{2x}$, and $\delta_y = \phi_{1y} - \phi_{2y}$. As a result of the adjoint operation, the interferogram I contains none of the time dependence or z-dependence that the optical waves themselves have. The part that has disappeared from the expression is the mean dynamic phase of the two arms.

It is a curious property of interferograms that, although they operate with intensities rather than amplitudes, they give a phase result proportional to ϕ, the phase of the electric fields themselves, and not to 2ϕ, the phase of the 2D oscillation wave.

The interferogram (4.25) consists of a constant C plus two cosine waves that oscillate with respect to the variable Δ. Thus, as we tune the OPD between the arms of the interferometer, the intensity modulates. Because we have a sum of two cosine waves, we can use the same approach that we did for the 1D wave composition formulas to calculate the phase of the interferogram peak with respect to Δ. Taking the derivative of I with respect to Δ, setting the derivative to zero, and solving for Δ, we obtain

$$\tan\Delta = \frac{A_{1x}A_{2x}\sin\delta_x - A_{1y}A_{2y}\sin\delta_y}{A_{1x}A_{2x}\cos\delta_x + A_{1y}A_{2y}\cos\delta_y}\,. \tag{4.26}$$

Although the result is no longer a propagating wave but rather a stationary interferogram, we find that the expression for the interferogram phase Δ is nearly the same as that of the geometric phase γ of the 1D propagating wave, (4.3). Thus, when the component amplitudes of the two arms are exactly matched, then the interferogram phase Δ exactly matches the propagating wave's geometric phase: $\tan\Delta = \tan\gamma$. Although it is not a requirement, this is easiest to see when the polarization states of the two arms are also exactly matched (i.e., $\delta_x = \delta_y$). In this special case, the interferogram expression (4.25) simplifies to

$$I = C + C\cos\delta\,. \tag{4.27}$$

When the amplitudes in the two arms differ, and the polarization states of the two arms are not the same, then attenuating one arm relative to the other will have an effect on the location of

the peak. In this case the wave geometric phase and the interferogram phase will not be the same. This is a feature that is not widely recognized in discussions about geometric phase, though it is an element in Pancharatnam's original paper.[8]

The results presented in (4.26) allow us to describe what happens in the case of "open loop" configurations of geometric phase measurements. Such open loop situations — where the polarization state output by the interferometer's sample arm is not matched to that in the reference arm — are the subject of lengthy discussion in the literature, but we now have a straightforward means of quantifying what happens in this case, and how this affects the geometric phase measurement. The closed loop configuration, when attenuation in the sample arm is negligible, will produce a phase value that agrees with geometric phase predictions, but if attenuation is significantly different in the two arms, especially when the two polarization states differ, then the interferogram phase will diverge from γ.

The fact is that the interferogram is not a wave in the same way that light is a wave. The interferogram is stationary, has different units, and is not a transverse wave. Keeping this in mind, it is only natural to observe that they will not in general possess the same phase, except under specific measurement conditions.

Finally, we can add that if the geometric phase γ is explicitly represented in the wave expression, such as by rewriting (4.24) in the form

$$\vec{E}_1 = e^{i(\psi_1+\gamma)} \begin{pmatrix} A_{1x}e^{-i\phi_{1x}} \\ A_{1y}e^{+i\phi_{1y}} \end{pmatrix}, \quad \vec{E}_2 = e^{i\psi_2} \begin{pmatrix} A_{2x}e^{-i\phi_{2x}} \\ A_{2y}e^{+i\phi_{2y}} \end{pmatrix}, \tag{4.28}$$

then we can see that the result (4.26) merely has the geometric phase added to it,

$$\tan(\Delta - \gamma) = \frac{A_{1x}A_{2x}\sin\delta_x - A_{1y}A_{2y}\sin\delta_y}{A_{1x}A_{2x}\cos\delta_x + A_{1y}A_{2y}\cos\delta_y}, \tag{4.29}$$

so that the geometric phase γ from the sample arm produces one shift, while term on the right hand side of the equation represents an additional phase shift produced by the interference of two beams of differing polarization state.

Although (4.29) looks equivalent to (4.3), it has a subtle difference. Since the interferogram "wave" is stationary, when we measure the wave we can detect shifts of the pattern with respect to a fixed reference plane. This allows us to measure phases wrapped within the range of $-\pi < \gamma \leq +\pi$. In the case of summing waves, however, there is no fixed reference plane available and so we can only measure the two wave peaks with respect to one another. This latter case constrains the geometric

phase values to the smaller range of $-\pi/2 < \gamma \leq +\pi/2$. This difference in geometric phase ranges can cause confusion when comparing our analysis with the existing literature, if an author does not make explicit whether he/she is referring to the electric field wave phase or to the interferogram phase.

The nonlinear behavior of the geometric phase has been suggested for use in fast optical switching.[33,34] Although this has only been discussed using 2D waves, the same behavior is already present in 1D. For the sum of two waves of different amplitudes, the location of the sum peak not only shifts closer to the location of the higher-amplitude wave peak, but this shift "accelerates" as δ approaches $\pm\pi/2$. At $\pm\pi/2$, the value of γ jumps by π as a result of the sudden change in reference location. Note, however, that this occurs at the point in which the two waves are closest to their destructive interference point, where the output intensity will be weak unless the two beams are substantially different in amplitude. For 2D waves, rather than being a location of minimum brightness, this is instead a location at which the fringe visibility will be minimum.[27,35] Thus, there is a tradeoff between measurement precision (using the light intensity or fringe contrast) and switching speed, since a large difference in amplitude between the two beams blunts the sharpness of the transition: with a larger difference in wave amplitudes, a larger change in δ is needed in order to effect a phase change.[36]

4.6 The novelty of the wave description of geometric phase

Whereas previous work in the literature has discussed geometric phase in the abstract formalism of matrix calculus or spherical trigonometry, we have shown that the same results can be derived entirely from considering the phases of waves and analyzing how the wave peak position changes when waves are added together. This makes it possible to visualize phase relationships from the waves themselves, with minimal mathematical abstraction.

Our approach provides an argument (Sec. 4.2) for why the unusual quantities of "instantaneous intensity"[37] (the square of the electric field, E^2) and "instantaneous Stokes vector" are used in the geometric phase literature. While previous researchers have used these without justification, other than that they work, we see that these arise naturally from calculating the position of the wave peak of a 2D polarized wave.

Whereas the existing literature calculates γ only before and after a homogeneous optical ele-

ment, it does not provide a model for calculating γ continuously as a wave propagates through a polarization element. The wave model that we present here shows how this is to be constructed.

We also found that the likely reason why so much of the geometric phase literature assumes closed-loop cycles is that this is a situation for which the interferogram phase and the wave phase are identical. If the reference arm polarization state and the sample arm state are different, then the interferogram will exhibit a phase shift in accordance with (4.26).

Finally, our results also provide a simple explanation for the differences between geometric phases of scalar waves and vector waves[31] — differences entirely due to the disparity between composing two parallel waves, (4.3), and composing two orthogonal waves, (4.10).

4.7 Appendix

4.7.1 Appendix A. Summing together N waves in 1D

The Harmonic Addition theorem states that the sum of N waves can itself be written as a wave[5]

$$\psi = \sum_i A_i \cos(kz - \omega t + \phi_i) = \bar{A}\cos(kz - \omega t - \bar{\phi})\,, \tag{4.30}$$

with an amplitude and phase of

$$\begin{aligned} \bar{A} &= \left[\sum_{i=1}^{N}\sum_{j=1}^{N} A_i A_j \cos(\phi_i - \phi_j)\right]^{1/2} \\ &= \left[\sum_{i=1}^{N} A_i^2 + 2\sum_{i=1}^{N}\sum_{j>i}^{N} A_i A_j \cos(\phi_i - \phi_j)\right]^{1/2}, \end{aligned} \tag{4.31}$$

$$\tan\bar{\phi} = \frac{\sum_{i=1}^{N} A_i \sin\phi_i}{\sum_{i=1}^{N} A_i \cos\phi_i}\,. \tag{4.32}$$

4.7.2 Appendix B. Comparing Eq. 13 to an expression in the existing literature

Equation 7 in Ref. 15, and Equation 1 in Ref. 16 give

$$\Phi = \arg\left\{\mu_\mathrm{p} + \mu_\mathrm{q} + (\mu_\mathrm{p} - \mu_\mathrm{q})\vec{s}_\mathrm{p}\cdot\vec{s}_\mathrm{a}\right\}. \tag{4.33}$$

From the definitions given in those papers, we can write $\mu_\mathrm{p} = e^{-i\delta}$, $\mu_\mathrm{q} = e^{i\delta}$. (Note that those references use the symbol δ to represent what we define as 2δ in (4.5).) Also, we can note that

$\vec{s}\cdot\vec{p}=\cos(2\psi)$ where ψ is an angle between the two points $\vec{s}_{\rm p}$ and $\vec{s}_{\rm a}$ on the Poincaré sphere. As a result, (4.33) becomes

$$\begin{aligned}\gamma &= \arg\left\{e^{-i\delta}+e^{i\delta}+(e^{-i\delta}-e^{i\delta})\cos(2\psi)\right\}\\ &= \arctan\left[\tan(\delta)\cos(2\psi)\right],\end{aligned} \tag{4.34}$$

where we have also replaced Φ with our geometric phase variable γ. In order to compare this expression with our own approach, we look to (4.11). Although we use A_1 and A_2 rather than the angle ψ, we can represent our amplitudes with $A_1=\cos\psi$, $A_2=\sin\psi$, so that (4.11) becomes

$$\tan\gamma=\tan(\delta)\,\frac{\cos^2\psi-\sin^2\psi}{\cos^2\psi+\sin^2\psi}=\tan(\delta)\cos(2\psi)\,, \tag{4.35}$$

which is in agreement with (4.34).

4.7.3 Appendix C. Hidden encoding of a global phase inside local phases

In order to see how a global phase can be encoded inside the apparently local phases defined in (4.24), we can consider the following example. In our Mach-Zender interferometer, we place a single retarder in the sample arm, with retardance δ. The light incident on the sample arm is linearly polarized at 0°(i.e. an H-wave) and the retarder is also oriented with azimuth 0°. Then the input and output states in the sample arm will be

$$E_1=e^{i\psi_1}\begin{pmatrix}1\\0\end{pmatrix}, \tag{4.36}$$

$$E_2=e^{i\psi_2}\begin{pmatrix}1\cdot e^{i\delta/2}\\0\cdot e^{-i\delta/2}\end{pmatrix}=e^{i[\psi_2+(\delta/2)]}\begin{pmatrix}1\\0\end{pmatrix}. \tag{4.37}$$

Thus, the polarization state is unchanged, but we have delayed the H-wave by half the retardance.

Next, we consider the case when the input polarization state is oriented at 45°(i.e., a D-wave), and the retarder fast-axis is also oriented with an azimuth angle of 45°. Thus, this setup is effectively equivalent to the previous case, but with our reference axis rotated. Still working in the x-y basis, the input and output waves of this case will be

$$E_1=e^{i\psi_1}\begin{pmatrix}\cos 45^\circ\\ \sin 45^\circ\end{pmatrix}, \qquad E_2=e^{i\psi_2}\begin{pmatrix}\cos 45^\circ e^{i\delta/2}\\ \sin 45^\circ e^{-i\delta/2}\end{pmatrix}. \tag{4.38}$$

The retardance δ in the output wave no longer takes on the appearance of a global phase, but is instead encoded within the local phases of the x and y components.

Chapter 5

Pancharatnam's discovery and why his geometric relationship resembles the addition of waves

The physical model presented in the previous chapter, which explains geometric phase as a shift experienced by a wave due to the relative displacement of its parts and their corresponding amplitudes, seems very different in nature to the geometrical relationship on Poincaré sphere found by Pancharatnam. If the physical model is accurate then there should be some equivalence or at least a connection to Pancharatnam's discovery.

This was my motivation to read again his paper with new understanding. Despite having read his paper multiple times, some things made more sense this time. After going through the paper I finally understood how Pancharatnam arrived to a concept that is equivalent to finding the peak of the sum of two waves.

Reading through Pancharatnam's paper I realized that there was a widespread misconception about his work. In his paper he focuses on interference and not on evolution of SOPs. I was then intrigued by the origin of the misconception. How is it possible that so many researchers cite Pancharatnam's work saying it is related to evolution of SOPs? This then motivated me to read about the historical events that surrounded the discovery of geometric phase, to confirm that Pancharatnam was not working on evolution of SOPs, and to track down the origin of the misconception. It turns out the misconception was actually introduced by analyzing Pancharatnam's work with the point of view of Berry's work. I clarify that the misconception is not severe and that Berry's work is correct. The misconception lies only in the interpretation of Pancharatnam's work. I start the chapter by giving background on the historical events related to the discovery of geometric phase and then I give a summarized and simplified explanation of Pancharatnam's paper.

Sivaramakrishnan Pancharatnam was born in Calcutta in 1934 in a family of remarkable scientists. He joined his brother S. Ramaseshan at Bangalore in the Raman Research Institute in 1953 when he was 19 years old.[38,39] At that time Ramaseshan's research supervisor was C. V. Raman, their uncle, and who received the Nobel prize in physics in 1930. Under C. V. Raman's instruction, Pancharatnam studied the behavior of absorbing bi-axial crystals, as he points out in the

opening of his now-famous paper, "Generalized theory of interference and its applications, part I. Coherent pencils".[8] This impressive work was published in 1956 when Pancharatnam was 22. In it, Pancharatnam proposes a definition for two beams of different states of polarization to be in phase: "The phase advance of one polarised beam over another (not necessarily in the same state of polarization) is the amount by which its phase must be retarded relative to the second, in order that the intensity resulting from their mutual interference may be maximum."[8] Berry named this *Pancharatnam's connection*.[9]

After providing expressions for the phase shift — what we now refer to as the geometric phase — Pancharatnam goes on to relate the expressions to the subtended solid angle on the Poincaré sphere. In Sections 5.1 & 5.2 below, we retrace Pancharatnam's derivation using a more modern approach that should be easier for modern readers to follow. Along the way, it becomes clear that Pancharatnam nowhere considers the case that is widely attributed to him, that the phase of polarized light changes after a cyclic evolution of its polarization.[9,21,40–45] This misconception of his actual work appears to be a result of the difficulty one encounters when reading Pancharatnam's paper, which often feels like it belongs to the 19th century, and also of confounding Pancharatnam's work with the closely related work of Berry.

Almost 30 years after Pancharatnam's original work, Michael Berry at the University of Bristol, unaware of Pancharatnam's work, discovered that an unexpected phase emerges after the adiabatic evolution of a quantum state around a cycle in parameter space.[10] In 1983, before his initial paper was published, Berry introduced the geometric phase to Barry Simon, who immediately coined it as *Berry's phase*.[46] By the end of 1986 Ramaseshan and Nityananda revived Pancharatnam's work and presented it as an example of Berry's phase.[47] Berry himself received Nityananda's manuscript and read it, but mentions that it was not until he visited Bangalore in July 1987 that he came to appreciate Pancharatnam's work. One can surmise that Nityananda's interpretation of Pancharatnam's phase was likely greatly influenced by Berry's personal approach to the geometric phase via cycles of states. This seems likely because Berry, on his return from India, prepared a new manuscript that revealed the connections between his and Pancharatnam's work, in which he explains Pancharatnam's work in terms of cycles of states.[48,49] Subsequent researchers have almost invariably followed this interpretation, with the result that Pancharatnam's actual achievement is obscured beneath the misconception.

As Berry himself pointed out, there have been several anticipations to geometric phase that arise

before Pancharatnam's work.[49] Vinitskii *et al.*, for example, mention the work of Rytov (1938) and Vladimirskii (1941) as precursors.[50] Oriol Arteaga has also pointed out that Fresnel and Arago in 1816 developed their "fifth" law of optical interference in such a way that a geometric phase term (at the time not well understood) had to be included in the interference equations.[45] However, as is commonly the case in science, every discovery can be traced to its anticipations, and we focus on Pancharatnam because his achievement is widely recognized by the scientific community.[51]

5.1 Pancharatnam's starting point

At the beginning of his manuscript, after explaining briefly some of the properties of the Poincaré sphere and Stokes parameters, Pancharatnam introduces the following theorem:

> *Theorem 1.* When [an electromagnetic] vibration of intensity I in the state of polarisation **C** is decomposed into two vibrations in the opposite states of polarisation **A** and **A′**, the intensities of the **A**-component and the **A′**-component are $I\cos^2(AC/2)$ and $I\sin^2(AC/2)$ respectively.

(See Fig. 5.1 for an illustration of the geometry.) Here Pancharatnam makes use of the fact that the angle $\angle A'C$, between points **A′** and **C**, subtended from the center of the Poincaré sphere, is complementary to angle $\angle AC$, and so writes both components in terms of $\angle AC$ alone. Since we will be making use of these angles in many of the equations below, we follow Pancharatnam and define the individual angles a, b, and c as

$$\left.\begin{aligned} c &= \angle AB \\ b &= \angle AC \\ a &= \angle BC \end{aligned}\right\} \tag{5.1}$$

Note that these angles a, b, and c on the left hand side of the equations have to be divided by two when used to describe the corresponding angles of the electric fields in Cartesian space.

Since this theorem is the basis for much of Pancharatnam's subsequent results, we show how one can derive it using a modern approach. Let us consider a monochromatic electromagnetic wave propagating along the z axis, i.e., $\mathbf{E} = \mathbf{E}_C \exp[i(kz - \omega t)]$. The electric field amplitude $\mathbf{E}_C$ can be written using elliptical polarization states $\{\hat{\mathbf{e}}_A, \hat{\mathbf{e}}_{A'}\}$ as a basis, with the properties $|\hat{\mathbf{e}}_{A,A'}| = 1$, and

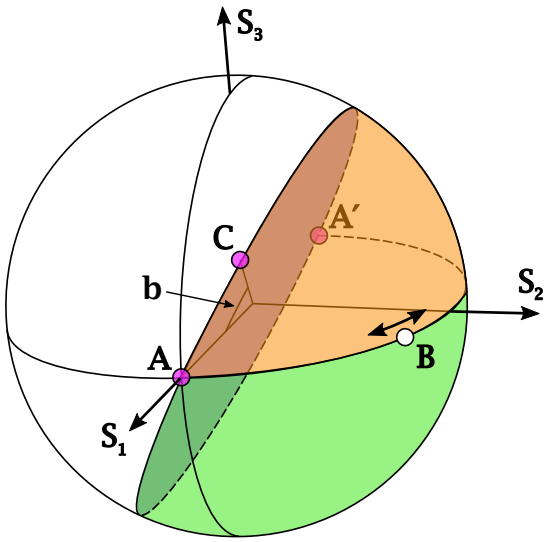

FIGURE 5.1. Polarization state **C** with respect to the two orthogonal states **A** and **A′**, represented on the Poincaré sphere. Point **B** represents an SOP lying between **A** and **A′**.

$\hat{\mathbf{e}}_A \cdot \hat{\mathbf{e}}^*_{A'} = 0$ (where $*$ represents the complex conjugate). Therefore,

$$\mathbf{E}_C = E\cos(\alpha)\hat{\mathbf{e}}_A + E\exp(i\beta)\sin(\alpha)\hat{\mathbf{e}}_{A'}, \tag{5.2}$$

where E is a real amplitude (thus, $I = E^2$), α controls the projection over the basis vectors, and β is the relative phase. With the above equation, we can represent any state of polarization (SOP) by modifying α and β.

For the common definition of the Stokes parameters, we would choose $\hat{\mathbf{e}}_A = \hat{x}$ and $\hat{\mathbf{e}}_{A'} = \hat{y}$, so that

$$\left.\begin{aligned} S_0 &= |E_x|^2 + |E_y|^2\,, \\ S_1 &= |E_x|^2 - |E_y|^2\,, \\ S_2 &= 2\mathrm{Re}[E_x E_y^*]\,, \\ S_3 &= 2\mathrm{Im}[E_x E_y^*]\,, \end{aligned}\right\} \tag{5.3}$$

However, this is only one specific choice. An equally valid choice is the following general form of the Stokes parameters[52]

$$\left.\begin{aligned} S_0 &= |E_A|^2 + |E_{A'}|^2\,, \\ S_1 &= |E_A|^2 - |E_{A'}|^2\,, \\ S_2 &= 2\mathrm{Re}[E_A E_{A'}^*]\,, \\ S_3 &= 2\mathrm{Im}[E_A E_{A'}^*]\,, \end{aligned}\right\} \tag{5.4}$$

where E_j, with $j \equiv A, A'$, are the components of $\mathbf{E}_C$, i.e., $E_j = \mathbf{E}_C \cdot \hat{\mathbf{e}}_j$. Let us consider **C** =

$\{S_1, S_2, S_3\}$, representing the SOP of $\mathbf{E}_C$, lies on the surface of the Poincaré sphere of radius S_0 with main axes $\{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3\}$. These definitions are illustrated in Fig. 5.1, where the $\mathbf{S}_i$ are with respect to a general elliptical basis and need not be with respect to the x-y basis. By using (5.2), the Stokes parameters for $\mathbf{E}_C$ are

$$\left.\begin{aligned} S_0 &= E^2 = I\,, \\ S_1 &= I\cos 2\alpha\,, \\ S_2 &= I\sin 2\alpha\cos\beta\,, \\ S_3 &= -I\sin 2\alpha\sin\beta\,. \end{aligned}\right\} \tag{5.5}$$

Here, 2α and β are the polar and azimuthal angles, respectively, in a spherical coordinate system with $\mathbf{S}_1$ being the polar axis, and $\{\mathbf{S}_2, \mathbf{S}_3\}$ the equatorial plane. Hence, if $\alpha = [0, \pi/2]$, and $\beta = [0, 2\pi]$, we can cover the entire surface of the sphere.

From (5.5), we see that the normalized Stokes vector of $\hat{\mathbf{e}}_A$ (shown in Fig. 5.1 as $\mathbf{A}$) is equal to $(1, 1, 0, 0)$. Therefore, 2α is the angle in Poincaré space that the SOP of $\mathbf{E}_C$ makes with the basis vector $\hat{\mathbf{e}}_A$ (cf. Fig. 5.1). Moreover, if we can define $\angle AC = 2\alpha$, we can use (5.2) to conclude that the intensity of the $\hat{\mathbf{e}}_A$-component is $I\cos^2(\angle AC/2)$ and the intensity of the $\hat{\mathbf{e}}_{A'}$-component is $I\sin^2(\angle AC/2)$, which is Pancharatnam's Theorem 1.

Pancharatnam next considers the problem of explaining interference between two non-orthogonal states of polarization. It was already well known that the interference of two beams of the same state of polarization (SOP) results in an intensity

$$I(I_1, I_2, \delta) = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\delta)\,, \tag{5.6}$$

where I_1 and I_2 are the individual intensities, and δ is the phase difference between the two beams. We can see that changing I_1 or I_2 only modifies the modulation contrast (visibility), but not the phase of the sum wave I.[17]

In order to see what happens when two beams of different SOP are combined, Pancharatnam considers the superposition of beams in states $\mathbf{A}$ and $\mathbf{B}$. He considers one of them, e.g., $\mathbf{A}$, and its orthogonal $\mathbf{A}'$, as the polarization basis, and finds the intensity of the sum by applying the following steps:

1. Adding the amplitude of the $\mathbf{A}$-component of $\mathbf{B}$ to the amplitude of $\mathbf{A}$, and squaring to give the intensity,

2. Adding the intensity of the $\mathbf{A}'$-component of $\mathbf{B}$ to the intensity obtained in step 1.

That is, the projection of $\mathbf{B}$ onto $\mathbf{A}$ gives the component of $\mathbf{B}$ that directly interferes with $\mathbf{A}$. The component of $\mathbf{B}$ orthogonal to $\mathbf{A}$ only adds a bias to the intensity.

Considering Pancharatnam's Theorem 1, we therefore decompose $\mathbf{B}$ into two beams of orthogonal polarizations, yielding the intensities

$$I_{B,A} = I_B \cos^2(c/2)\,, \tag{5.7}$$

$$I_{B,A'} = I_B \sin^2(c/2)\,, \tag{5.8}$$

where I_B is the intensity of $\mathbf{B}$, and c is the angle between the two states on the Poincaré sphere, and is thus twice the angle between the states in Cartesian space. For step 1, we combine I_A (the intensity of $\mathbf{A}$) with $I_{B,A}$ using (5.6), giving

$$\begin{aligned} &I(I_A, I_{B,A}, \delta) = \\ &\quad I_A + I_B \cos^2(c/2) + 2\big[I_A I_B \cos^2(c/2)\big]^{1/2} \cos(\delta)\,. \end{aligned} \tag{5.9}$$

This expression represents the intensity that results from the addition of the $\mathbf{A}$-component of both beams.

Finally, with step 2, we add $I_{B,A'}$ to obtain the intensity

$$\begin{aligned} I = I_A &+ I_B \cos^2(c/2) + I_B \sin^2(c/2) \\ &+ 2\big[I_A I_B \cos^2(c/2)\big]^{1/2} \cos(\delta)\,, \end{aligned} \tag{5.10}$$

which simplifies to

$$I = I_A + I_B + 2\sqrt{I_A I_B}\cos(c/2)\cos(\delta). \tag{5.11}$$

The above equation is Pancharatnam's Eq. (1) in Ref. 8. This expression has the advantage that the factor containing the phase δ between the beams and $\cos(c/2)$ — the square root of what Pancharatnam calls their "similarity factor" — are separated as a product.

While δ is the phase delay between the $\mathbf{A}$-component of both beams, Pancharatnam gives a couple of arguments to justify that this phase is equal to the phase between $\mathbf{A}$ and $\mathbf{B}$, supported by (5.11). This equation shows that the intensity I of the beams' superposition varies sinusoidally, with δ as the phase of I. This phase is independent of what polarization states are involved, and is not even dependent on the relative similarity of the two states. Rather, as the two states become more

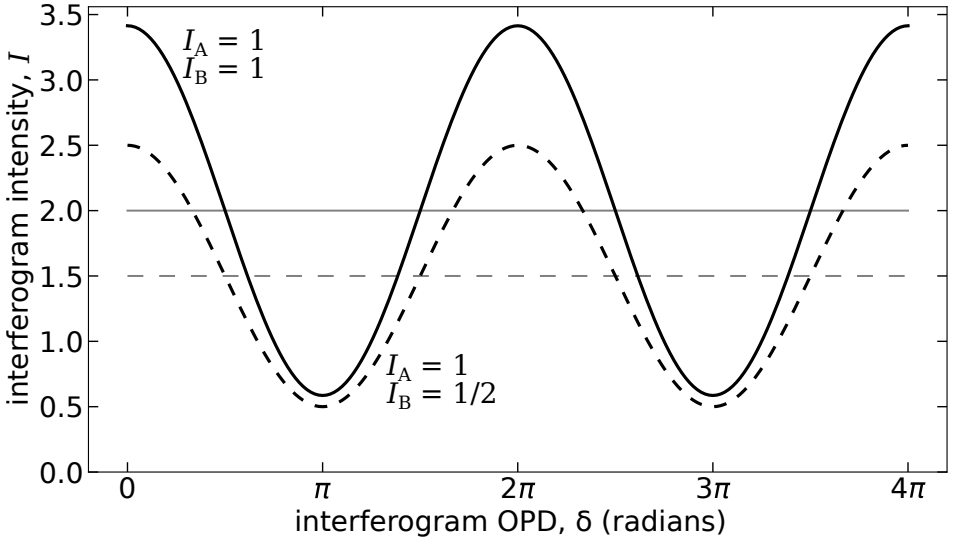

FIGURE 5.2. The "original wave" interferogram is obtained in (5.11) when the two intensities I_A and I_B are equal and held fixed as we vary the relative phase δ between them. The "attenuated wave" occurs when we reduce the intensity of one beam, and again vary the phase between the two beams.

"dissimilar", the modulation contrast (as illustrated in Fig. 5.2) decreases, but the phase remains unchanged. Second, if either of the two intensities I_A or I_B changes, this will modify I but does not change the value of δ. We can see this in the interference of two beams, regardless of their states of polarization: if we modify only the intensities of the input beams, we do not see a displacement of interference fringes. These results are what Pancharatnam has in mind when he titles his paper the "generalized theory of interference".

To compare with more modern approaches, we can analyze the same case using the Jones calculus. In order to simplify the math, we will simply define the state of the first beam to be horizontal, while that of the second beam is oriented at an angle of θ with respect to the horizontal. We also allow there to be a propagation phase delay δ between the two beams. Therefore, we have the two electric field vectors

$$\vec{E}_A = |\vec{E}_A|\hat{x}, \qquad \text{and} \qquad \vec{E}_B = |\vec{E}_B|e^{-i\delta}\left(\cos\theta\hat{x} + \sin\theta\hat{y}\right). \tag{5.12}$$

The amplitude of their sum is

$$\vec{E}_A + \vec{E}_B = \left(|\vec{E}_A| + |\vec{E}_B|\cos\theta\, e^{-i\delta}\right)\hat{x} + |\vec{E}_B|\sin\theta\, e^{-i\delta}\hat{y}, \tag{5.13}$$

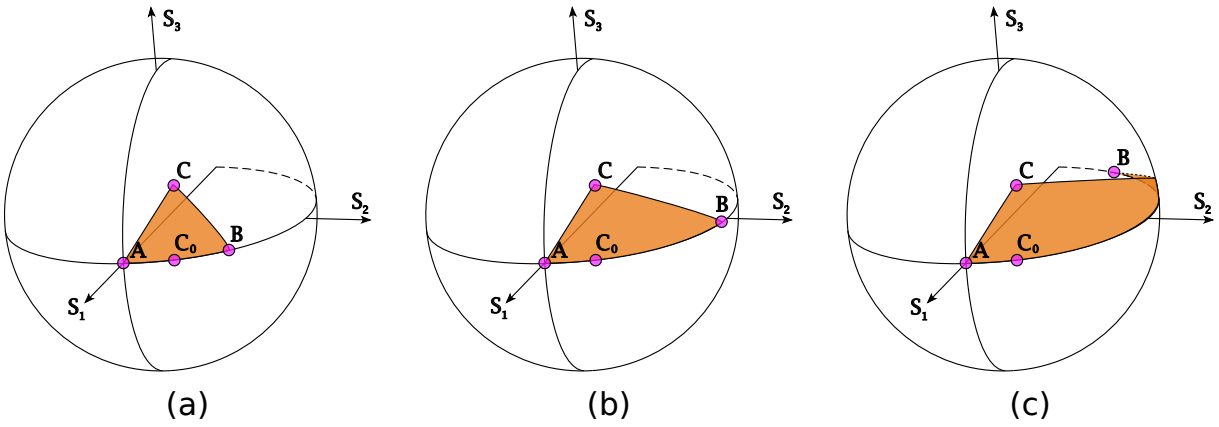

FIGURE 5.3. The increasing area (drawn in orange) on the surface of the sphere, formed by the spherical triangle between points **A**, **B**, and the result of their addition, **C**. When beams in states of polarization corresponding to **A** and **B** are added in phase the resulting polarization corresponds to $\mathbf{C}_0$, lying in the geodesic between them. When **B** approaches $\mathbf{A}'$, then antipode of $\vec{A}$, then the orange area becomes that of Fig. 5.4.

so that intensity of the two beams' interference is

$$\begin{aligned} I &= \left(\vec{E}_A + \vec{E}_B\right) \cdot \left(\vec{E}_A + \vec{E}_B\right)^* \\ &= \left(|\vec{E}_A| + |\vec{E}_B| \cos\theta\, e^{+i\delta}\right)\left(|\vec{E}_A| + |\vec{E}_B| \cos\theta\, e^{-i\delta}\right) \\ &\quad + \left(|\vec{E}_B| \sin\theta\, e^{+i\delta}\right)\left(|\vec{E}_B| \sin\theta\, e^{-i\delta}\right) \\ &= |\vec{E}_A|^2 + |\vec{E}_B|^2 + 2|\vec{E}_A||\vec{E}_B| \cos\theta \cos\delta\,, \end{aligned} \tag{5.14}$$

which is exactly Pancharatnam's expression [given as (5.11) above], since $c = 2\theta$.

This can be generalized to an arbitrary pair of elliptical states, considering two waves with electric field amplitudes $\mathbf{E}_A$ and $\mathbf{E}_B$, and a phase difference δ. The intensity of their superposition is

$$\begin{aligned} I &= (\mathbf{E}_A + e^{i\delta}\mathbf{E}_B) \cdot (\mathbf{E}_A^* + e^{-i\delta}\mathbf{E}_B^*) \\ &= |\mathbf{E}_A|^2 + |\mathbf{E}_B|^2 + 2\mathrm{Re}\left\{e^{-i\delta}(\mathbf{E}_A \cdot \mathbf{E}_B^*)\right\}. \end{aligned} \tag{5.15}$$

The result of $\mathbf{E}_A \cdot \mathbf{E}_B^*$ can be inferred from the discussion about Pancharatnam's Theorem 1, which gives $E_A E_B \cos(c/2)$, where E_A, E_B are the real amplitudes of the electric fields. Therefore, we conclude that

$$\begin{aligned} I &= |\mathbf{E}_A|^2 + |\mathbf{E}_B|^2 + 2E_A E_B \cos(c/2)\cos(\delta) \\ &= I_A + I_B + 2\sqrt{I_A I_B}\cos(c/2)\cos(\delta), \end{aligned} \tag{5.16}$$

which agrees with Pancharatnam's result.

5.2 The relationship between intensities and locations on the Poincaré sphere

Following his expression for the interference of nonorthogonal beams, Pancharatnam requires only a few short steps (Sec. 4 of his paper) to develop his famous solid angle formula for the geometric phase. Here he considers the decomposition of a beam of state of polarization **C** into two nonorthogonal beams of polarization states **A** and **B**. If (5.11) is a general equation for the interference of two beams with different SOPs, then one should be able to re-arrange the equation to solve for the phase delay δ between them via the intensities of **A** and **B** (but which Pancharatnam labels I_1 and I_2):

$$\cos(\delta) = \frac{I - (I_A + I_B)}{2\sqrt{I_A I_B}\cos(c/2)}\,. \tag{5.17}$$

We can interpret this result as saying that "whatever this phase δ may be, it maintains a specific relationship between the intensities of the beam being decomposed (**C**) and the intensities of the beams that result from the decomposition (**A** and **B**)."

While he could have finished with this numerical formula for δ, he went a step further and realized that this equation expresses a solid angle relationship between the SOPs of the three beams. To explain this he uses electric field vectors $\vec{E}_A$ and $\vec{E}_B$ as components of a sum vector $\vec{E}_C = \vec{E}_A + \vec{E}_B$. From vector analysis, it is easy to see that the part of $\vec{E}_B$ perpendicular to $\vec{E}_A$ also has to be equal to the part of $\vec{E}_C$ perpendicular to $\vec{E}_A$. Writing this in terms of the angles a, b, and c from (5.1), we have

$$\sqrt{I_B}\sin(c/2) = \sqrt{I}\sin(b/2)\,. \tag{5.18}$$

Taking the square of the components to find the intensities results in the proportions

$$I_A = I\,\frac{\sin^2(a/2)}{\sin^2(c/2)} \qquad \text{and} \qquad I_B = I\,\frac{\sin^2(b/2)}{\sin^2(c/2)}\,. \tag{5.19}$$

By expressing the intensities I_A and I_B in terms of the total intensity I, together with the angles between states on the Poincaré sphere, we can rewrite (5.17) as

$$\cos(\delta) = \frac{\sin^2(c/2) - \sin^2(a/2) - \sin^2(b/2)}{2\sin(a/2)\sin(b/2)\cos(c/2)}\,, \tag{5.20}$$

which is Pancharatnam's Eq. 4. Now the equation is entirely in terms of the angles between states, rather than their intensities, and he recognizes that the form of this expression is almost the same

as that of a solid angle formula. In fact, if we re-express the above relationships in terms of states **A**, **B**, and **C**′ rather than **A**, **B**, and **C**, where **C**′ is the antipodal point to **C** on the Poincaré sphere, we obtain[53]

$$\cos(\delta) = \frac{1 - \cos^2(c/2) - \cos^2(a'/2) - \cos^2(b'/2)}{2\cos(a'/2)\cos(b'/2)\cos(c/2)}\,, \tag{5.21}$$

which does have the recognizable form of a solid angle formula. The primes indicate that the angles are to be taken with respect to **C**′ rather than **C**. That is, $b' = \angle AC'$, and $a' = \angle BC'$. Making use of the solid angle formula, we obtain[54]

$$\delta = \pi - \Omega'/2\,, \tag{5.22}$$

where Ω' is the angle subtended by the spherical triangle $\mathbf{A}, \mathbf{B}, \mathbf{C}'$ from the center of the sphere. (Pancharatnam actually expresses this as $\delta = \pi - \frac{1}{2}E'$, where E' is the spherical excess of the triangle, and therefore also equal to the solid angle $\Omega' = E'$.) The sign of δ given here corresponds to describing the sequence of states $\mathbf{A} \to \mathbf{B} \to \mathbf{C}'$ in a clockwise sense. If the direction of the sequence is reversed, then the sign flips.

These last statements, in which Pancharatnam works out the correct sign of the solid angle, is the only location in the paper where he talks about a sequence of states. However, it is clear from context that he is not referring to a cycle of polarization states but rather to the phase relationships between the two output states **A** and **B**, and the state **C** from which they were decomposed.

5.3 Extending the reasoning to the orthogonal case

Considering that Pancharatnam found the phase between two beams using intensity of interference, his definition does not apply for the case in which states **A** and **B** are orthogonal. However, he provides a geometric argument to show that the formula can still be applied in the limit as the states approach orthogonality. Figure 6.4a shows an initial situation with states of polarization **A** and **B**, and the state **C** obtained by their sum. This is therefore the inverse of the case treated in Sec. 5.2 above, where we split a beam in two; this time we are adding **A** and **B** with a phase δ between them such that (5.11) in Sec. 5.1 holds. A point $\mathbf{C}_0$ lying on the geodesic arc **AB** describes a state of polarization that results from adding the beams of polarization **A** and **B** with no phase between them. Recalling that the solid angle used by Pancharatnam is the one subtended by triangle **ABC**′ rather than **ABC**, this situation gives a solid angle of $\Omega' = 2\pi$. That is, when $\mathbf{C}_0$ lies on the smaller

geodesic between **A** and **B**, the lune enclosed between geodesic arcs $\mathbf{AC'B}$ and $\mathbf{AC_0B}$ spans a solid angle of half the sphere, while the lune enclosed within **ACB** and $\mathbf{AC_0B}$ spans a solid angle of zero.

Next we modify the polarization state of **B** so that it moves further away from **A** along the equator of the Poincaré sphere, as shown in Fig. 6.4b, and then Fig. 6.4c. As **B** moves away from **A**, the enclosed solid angle subtended by **ABC** increases. As **B** approaches the point orthogonal to **A** (this point is labelled by Pancharatnam as $\mathbf{A'}$), the geodesic **ACB** becomes half a great circle, as in Fig. 5.4. One might argue that in this case the enclosed area becomes undefined, since we can draw the geodesic connecting **A** and **B** in either a clockwise or an anticlockwise sense. However, if we note that the geodesic from **A** to **B** has until this point always passed through the intermediate point $\mathbf{C}_0$, then it will for this limit case as well. The solid angle that relates the phase between the two beams is therefore the one enclosed between the two geodesic arcs $\mathbf{AC_0A'}$ and $\mathbf{AC'A'}$ (where $\mathbf{A'}$ coincides with the point written as **B** in Fig. 5.4). This is a spherical lune (drawn in green in the figure) whose solid angle subtended from the origin is exactly twice the value of the angle α formed between states $\mathbf{C}_0$, **A**, and $\mathbf{C'}$ at the surface of the sphere.

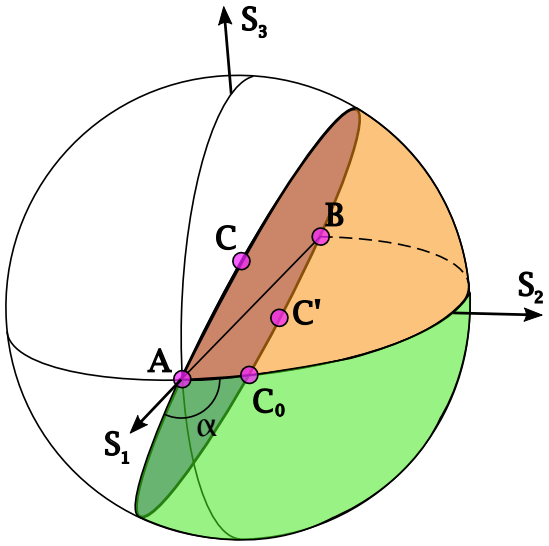

FIGURE 5.4. If we continue the evolution shown in Fig. 6.4 until state **B** becomes orthogonal to **A**, then we form a spherical lune (drawn in orange) between geodesic arcs $\mathbf{AC_0B}$ and **ACB**. The complementary spherical lune, drawn in green, between geodesic arcs $\mathbf{AC_0B}$ and $\mathbf{AC'B}$ corresponds to the solid angle Ω giving the geometric phase δ between **A** and **B** needed to obtain **C** upon addition.

Using (5.22) finally allows us to equate the phase difference between beams **A** and **B** to the angle α: $\delta = \alpha$. That is, the angle α (or half the solid angle Ω) is equal to the phase δ that one

must retard state $\mathbf{A}'$ from $\mathbf{A}$ in order that both be correctly decomposed from input state $\mathbf{C}$.

5.4 Interference of the components transmitted by an analyzer

It was of practical importance for Pancharatnam to have an expression for the phase of two arbitrarily-polarized beams transmitted through an analyzer because this was the configuration has was using to measure the pattern transmitted by biaxial crystals such as iolite. Finding such an expression is the purpose of Section 8 of Pancharatnam's manuscript.

Pancharatnam first defines $\mathbf{D}$ as the state of polarization transmitted by the analyzer. Unfortunately, Pancharatnam once again reuses the symbol $\mathbf{C}$ here. We write this as $\mathbf{D}$ to avoid confusion with the states defined in the earlier sections. He also uses I_1 and I_2 in place of I_A and I_B, but we have retained the latter for consistency. Pancharatnam also reuses the symbols a, b, and c for the angles between states, but since the definition of $\mathbf{C}$ has changed, we instead define θ_{ij} to represent the angle between points i and j on the surface of the Poincaré sphere, as subtended from the center of the sphere. His goal is to relate the phase of the interference transmitted by the analyzer to the phase difference δ between the two input states (see Fig. 5.5).

When beams $\mathbf{A}$ and $\mathbf{B}$ are incident on the analyzer, the transmitted intensity I_D will be the component of state $\mathbf{A}$ along direction $\mathbf{D}$, plus the component of state $\mathbf{B}$ also along $\mathbf{D}$, while incorporating an unknown phase difference δ' between $\mathbf{A}$ and $\mathbf{B}$:

$$\begin{aligned} I_D = {} & I_A \cos^2(\theta_{AD}/2) + I_B \cos^2(\theta_{BD}/2) \\ & + \sqrt{I_A I_B} \cos(\theta_{AD}/2) \cos(\theta_{BD}/2) \cos(\delta') \,. \end{aligned} \tag{5.23}$$

If we use the analyzer oriented at angle $\mathbf{D}'$ orthogonal to $\mathbf{D}$, then we would get a different intensity $I_{D'}$ and an unknown phase difference δ'':

$$\begin{aligned} I_{D'} = {} & I_A \sin^2(\theta_{AD'}/2) + I_B \sin^2(\theta_{BD'}/2) \\ & + \sqrt{I_A I_B} \sin(\theta_{AD'}/2) \sin(\theta_{BD'}/2) \cos(\delta'') \,. \end{aligned} \tag{5.24}$$

From Fig. 5.5, we can see that the area enclosed by $\mathbf{DAD'BD}$ is a spherical lune. The phase δ' needed to generate state $\mathbf{D}$ from adding $\mathbf{A}$ and $\mathbf{B}$ is given by $\delta' = \pi - \frac{1}{2}\Omega'$, where Ω' is the solid angle subtended by the spherical triangle $\mathbf{ABD'}$ (drawn in green in the figure). In a similar fashion, the phase δ'' needed when adding $\mathbf{A}$ and $\mathbf{B}$ to get $\mathbf{D}'$ is given by $\delta'' = \pi - \frac{1}{2}\Omega''$, where Ω'' is the solid angle of $\mathbf{ABD}$ (drawn in orange).

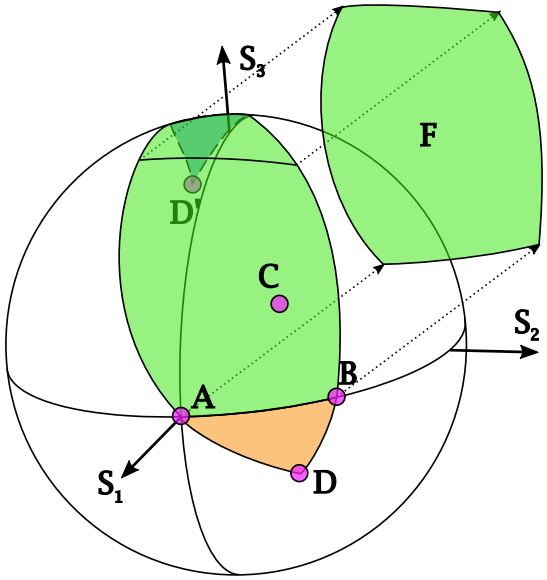

FIGURE 5.5. The spherical surface elements used to calculate the intensity transmitted by an arbitrary analyzer (state **D**) when two states **A** and **B** are incident upon it.

Since both δ' and δ'' are related to corresponding solid angles, if we subtract the two, we obtain the solid angle subtended by quadrangular area F in the sphere[21,55]

$$F = \pm(\delta'' - \delta') \,. \tag{5.25}$$

Confusingly, Pancharatnam yet again reuses the symbol C to indicate this quadrangular area. We will instead use F, as Fig. 5.5 does. The sign of F is determined by whether the sequence of points **ABD** proceeds in a clockwise or anticlockwise fashion. Pancharatnam then takes both I_D and $I_{D'}$ and adds them together to get the total intensity, equal to the intensity of **C** that depends on δ but now expressed in terms of δ' and F:

$$\begin{aligned} I = I_A + I_B + 2\sqrt{I_A I_B}\big[&\cos(\theta_{AD}/2)\cos(\theta_{BD}/2)\cos(\delta') \\ &+ \sin(\theta_{AD'}/2)\sin(\theta_{BD'}/2)\cos(\delta' \pm F)\big] \,. \end{aligned} \tag{5.26}$$

Now that we have the same angle δ' in both terms inside the square brackets, we can recognize that this is a standard expression for the spherical excess of a triangle, so that the equation simplifies to[54]

$$I = I_A + I_B + 2\sqrt{I_A I_B}\cos(c/2)\cos(\delta' + \tfrac{1}{2}\Omega) \,. \tag{5.27}$$

This equation is now identical in form to (5.11), from which we can then say that $\delta' = \delta - \frac{1}{2}\Omega$. Now we have a means of calculating δ' for a given pair of input states **A** and **B**, together with the

analyzer orientation giving **D**. Using this in (5.23), we can now calculate the intensity transmitted by the analyzer depending on the original phase difference δ between the two input beams.

5.5 Addressing the misconception

Through an impressive set of spherical trigonometry manipulations on the Poincaré sphere, Pancharatnam found that polarization states have specific phase relationships that are generally not taken into account, unless one is performing interferometric measurements. This was the case for him, since he was analyzing the light transmitted through dichroic biaxial crystals. The correct analysis of these measurements required that he incorporate this new phase, what we now refer to as the geometric phase, into his equations.

Among his less-known results is Equation 5.17, which shows that if one knows the intensities of two input beams as well as the intensity of their interference, one can infer the phase difference δ between the two input beams from these simple intensity measurements alone. Also notable is (5.27), which demonstrates that the phase transmitted by an analyzer is not in general equal to the phase of a wave incident upon it. While it might seem that Sec. 5.4 is not essential, it makes an important point that is often overlooked: the interferogram phase will not in general be the same as the phase difference δ between the two beams. If we want to measure the phase difference δ, we must pass the addition of the beams through an analyzer to obtain the interferogram. We can use (5.27), however, to relate the measured phase to δ.

The recently published "wave description of geometric phase" interprets the geometric phase as a shift in the wave peak location away from the midpoint between the peaks of the two input waves.[56] One can see the close connection between the wave description and Pancharatnam's approach by considering (5.11) (Pancharatnam's Eq. 1), which expresses the intensity produced by adding two waves as being a bias value $(I_1 + I_2)$ plus a single cosine term with amplitude $\sqrt{I_1 I_2}\cos(c/2)$. Thus, Pancharatnam is also considering the case of two cosine waves summing together into a single cosine output wave.

In (5.14), we were also able to show how Pancharatnam's results are derived using the familiar modern approach, not available to Pancharatnam, of the Jones calculus. The Jones calculus explicitly forms a pair of 2D electric field vectors, and adding these two waves produces a cosine factor $\cos(\delta)$ that includes the same phase delay that Pancharatnam obtained, but which does not seem

to have been replicated until Michael Berry's work in 1987.[9]

Berry naturally interpreted Pancharatnam's work through the lens of his own work, which approached the geometric phase as a shift resulting from the evolution of a quantum state around a cycle in parameter space. Upon reading Pancharatnam's approach, he saw that this could easily be a cycle of polarization states generating this phase shift. Pancharatnam, however, was focused on interferometric measurement and not on modeling the evolution of polarization states. As it turns out, the two are equivalent, and so the misunderstanding is not at all a serious one. In the literature, many authors actually refer to this shift due to polarization state evolution as a "Pancharatnam-Berry phase". This seems the ideal choice, since "Pancharatnam phase" should perhaps be more narrowly defined as a shift resulting from adding polarized waves.

Chapter 6

Differences between geometric phase and propagation phase: clarifying the boundedness problem

The next goal of this thesis is using the physical model developed in Chapter 4 to understand features of geometric phase theory that are still not widely understood, such as whether geometric and propagation phase can be simply added, and whether the geometric phase is unbounded (i.e. can take on any values without limit) or bounded (i.e. limited to values between $-\pi$ and $+\pi$). We show experiment results that improve on previously published work, in that the differences between geometric and propagation phase, and the meaning of geometric phase bounds, are shown clearly. We restrict ourselves to discussing Pancharatnam-Berry phase — the geometric phases induced by either transformations of the polarization state,[57] or by changes in phase induced by multiple-beam interference.[56] Thus, our discussion does not consider the case of "spin-redirection phase" (Aharonov-Anandan phase), in which the propagation vector of a beam traces a solid angle on the Bloch sphere.[58]

In the discussion below, we first summarize previous experiments exploring the boundedness problem, done by Bhandari, Hariharan, and their collaborators. A casual reading of this literature would lead one puzzled. Why would such respected authors conclude their work with conclusions that feel so arbitrary? We explain why this arbitrariness is only apparent, and that it is simply a result of leaving some of their evidence and conclusions unstated. We try to fill in those gaps.

Historical methods for calculating geometric phase γ have most often been based on paths on the Poincaré sphere. In order to link the historical discussion with a physical model for γ — the wave description of geometric phase — Sec. 4.3 reviews how to use the Jones calculus to calculate geometric phase for a simple retarder, pair of retarders, a retarder triplet (a HWP sandwiched between two QWPs, often abbreviated as a "QHQ"), or any system of retarders possessing a homogeneous Jones matrix.[59]

In Sec. 6.3, we introduce our own experiments to observe geometric and propagation phase in a white light Mach-Zehnder interferometer. By adjusting either the QHQ or the fold mirror position, we tune the geometric phase or the propagation phase in the measurement arm of the interferometer,

while leaving the reference arm stationary. The resulting interferograms clearly show the behavioral differences of the two types of phase. That is, the coherence envelope is unaffected by changes in the geometric phase, whereas changes in propagation phase cause a displacement of the coherence envelope. In Sec. 6.4, we switch from a Mach-Zehnder to a Michelson interferometer and repeat the same experiments using a line-imaging spectrometer. The motion of the spectral fringes here too shows a cyclical behavior produced by geometric phase shifts, but noncyclical (continuous shift) behavior produced by propagation phase shifts.

On the basis of these experimental results, we explain how one can argue for γ being bounded or unbounded, depending on one's choice of convention. For some experiments, considering γ to be bounded produces a simpler, or a more useful, interpretation of results. For other experiments, the unbounded case provides the simpler viewpoint.

6.1 Previous experiments on the boundedness problem

The first experiments that we are aware of that directly address the boundedness problem for geometric phase are the two reported by Bhandari in 1988.[1,2] For both experiments, Bhandari uses the setup shown in Fig. 6.1, employing a two-frequency stabilized single-mode He-Ne Zeeman laser as the light source in a Mach-Zender interferometer. Because the two Zeeman laser wavelengths have long coherence lengths, compensation for optical path difference (OPD) in the interferometer is not needed. Since the two wavelengths are orthogonally polarized, they are split by the polarizing beamsplitter (PBS), with one becoming the reference arm, while the second beam passes through optical elements before recombining at the final beamsplitter. The measurement arm beam starts horizontally polarized (**H**), is converted by the QWP to right-circularly polarized (**R**), then by the HWP to left-circularly polarized (**L**), and is projected onto a vertically polarized state (**V**) by the linear polarizer (LP). The resulting beam then combines with the reference arm beam at the final beamsplitter. Rotating the HWP by an angle θ thus produces a geometric phase $\gamma = 2\theta$ in the measurement arm. The detector measures the phase difference in the two arms via the intensity of the beat frequency between the two beams.[60]

Since the rotation of the HWP is continuous, one can watch the behavior of the interferogram over multiple rotations. Figure 6.2a shows a simulated set of interferograms to help visualize Bhandari's measurement, in which the abscissa of each interferogram is path difference Δ between

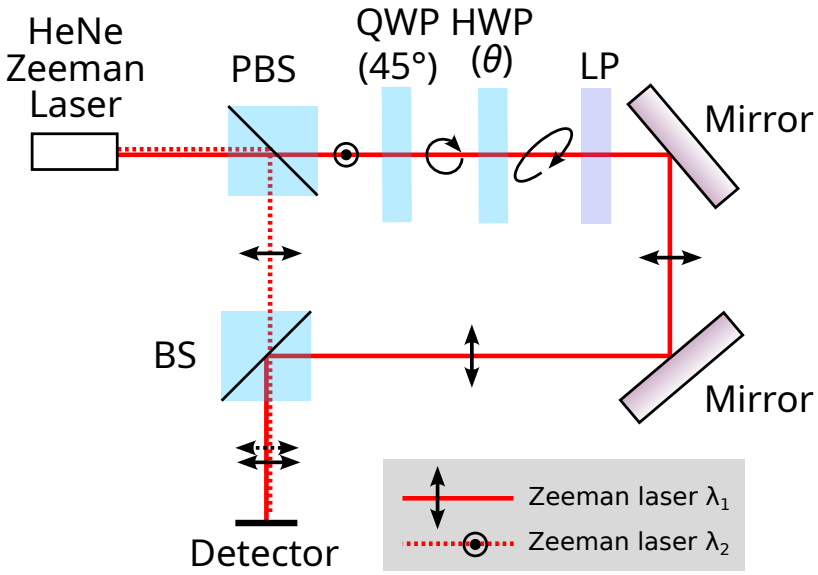

FIGURE 6.1. The experiment setup used by Bhandari in Ref 1. BS: beamsplitter, PBS: polarizing beamsplitter, LP: linear polarizer.

the two arms. The figure shows $2\pi/5$ steps of increasing phase, different from those used by Bhandari but simpler for visualization, and which correspond to $\theta = \pi/5$ rotation of the HWP. Figure 6.2b replicates Bhandari's figure showing the change in geometric phase with rotation of the HWP.

From a theoretical standpoint, rotating a HWP by a half-turn returns it to the same configuration, so that the model of any system using such a HWP should be periodic with 180°periods. However, if we track the position of the measured interferogram fringes, as indicated by the blue arrow in Fig. 6.2, then we see a continuous motion as the HWP rotates through 180°, 360°, and more. This continuous motion is what led Bhandari to declare the geometric phase to be unbounded, despite the discordance with needing periodicity. While rotating the HWP through two complete rotations, an experiment can watch as the interferogram fringe undergoes a steady shift by four full cycles, as we succeeded in replicating this result in our own experiments (Fig. 6.6 below). However, the interpretation of this result is not a straightforward as it may seem.

A second research group, involving Hariharan, Roy, & Larkin, further explored the boundedness problem with white light interferometry experiments reported in 1992 and 1994.[3,4] In these experiments, Hariharan and his co-authors used a Sagnac interferometer to avoid introducing propagation phase. Figure 6.3 shows the corresponding system optical layout, with either a HeNe laser or a white light source for the illumination. The "set of retarders" labelled in the figure corresponds to the following sequence of components: an achromatic circular polarizer (CP), followed by

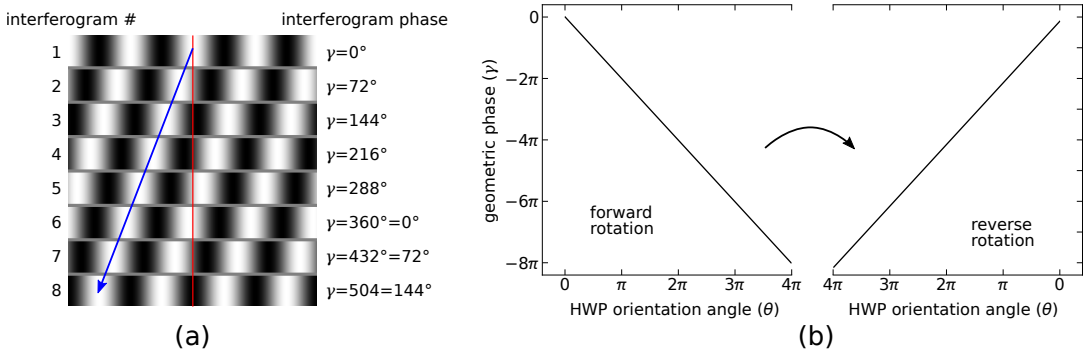

FIGURE 6.2. (a) A simulated set of interferograms to replicate the data used by Bhandari for generating (b), which shows the geometric phase measurements from the interferograms while rotating the HWP by two full rotations forward, and then two rotations back to the starting position (see Fig. 3 of Ref. 2). The red vertical line in (a) indicates the 0° reference position on the detector, while the blue arrow indicates the direction of motion of the interferogram fringes.

a QHQ, followed by a second achromatic circular polarizer. After the input, light passes through a linear polarizer oriented at 45°with respect to the plane of incidence of the interferometer and split equally into two component beams by a polarizing beamsplitter. Since the two beams traveling in opposite directions have orthogonal states of polarization (SOPs), they experience equal amounts of geometric phase shift but with opposite sign, and the propagation phase for both beams is identical.

As in Bhandari's experiments, Hariharan *et al.* found that rotating the achromatic HWP produces a continuous movement of the interferogram fringes. This remains true when operating with white light, but one can also observe that, unlike for the case of propagation phase, the coherence envelope in this case remains remains fixed. As a result Hariharan & Larkin conclude their paper with the perplexing statement that "the geometric phase, even though it is unbounded, can only be measured to modulo 2π." While this conclusion is entirely true, the authors' lack of elaboration on that statement can cause confusion. Similar statements can be found elsewhere, with similar potential to cause misunderstanding:

> While the Pancharatnam phase is additive, it does not affect the optical path difference.[61]

Similarly, Gordon Love, a co-worker of Bhandari, summarizes the research on the boundedness problem with the following:

> Do Pancharatnam phase shifts greater than 2π have a physical significance? This essentially philosophical question depends on one's definition of 'physical significance' [and depends] on one's perspective.[62]

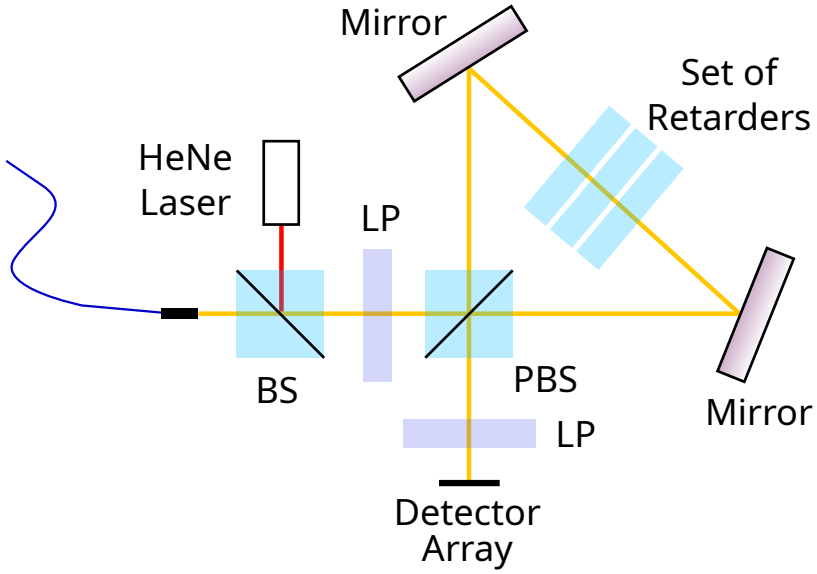

FIGURE 6.3. The experiment setup used by Hariharan *et al.* in Refs 3 & 4. The "set of retarders" is a combination of two circular polarizers (CP) sandwiching a QHQ: CP-QHQ-CP. PBS: polarizing beamsplitter, LP: linear polarizer. Both linear polarizers are oriented at azimuth angles of 45°.

6.2 Geometric phase due to a retarder or a QHQ

The recently developed wave description of geometric phase provides a physical model for the geometric phase that gives more insight into the behavior of the geometric phase than calculation methods alone can achieve.[56] We aim to use this physical model to resolve the boundedness problem, and so we first review how this wave description deals with propagation through optical systems containing sequences of waveplates.

As a wave propagates through a sequence of polarization components, the geometric phase of the wave can be calculated using Jones calculus.[59] We need only pay close attention to what happens to the two components of the polarization vector, and at the final step of the analysis, apply the geometric phase equation:[56]

$$\tan(\gamma + n\pi) = \tan\left(\frac{\delta}{2}\right)\frac{A_x^2 - A_y^2}{A_x^2 + A_y^2}, \tag{6.1}$$

where we have assumed the reference plane for calculating γ to lie at the midpoint between the peaks of the E_x and E_y component waves of the Jones vector. Here, A_x and A_y are the amplitudes, and δ the relative phase difference, of the x and y electric field components. Although (6.1) is obtained from Eq. 11 in Ref. 56, we have generalized the expression to allow solutions of the form $\gamma + n\pi$ instead of simply γ, in order to facilitate the coming discussion of phase bounds.

First we consider the simple case of a 45°linear polarized beam passing through a QWP with a

fast axis azimuth angle of 0°. Upon entering the retarder, the beam will split equally into horizontal and vertical components of the electric field, such that (6.1) gives $\gamma = 0$. Likewise, any polarization state whose two components are split equally upon entering the retarder will likewise produce $\gamma = 0$. The collection of these equal-split states forms a great circle on the Poincaré sphere — a circle whose axis of rotation passes through the two eigenstates of the retarder. In this case, the axis passes through $s_1 = 1$ and $s_1 = -1$, corresponding to linear polarizations of 0°and 90°.

Note, however that an elliptically polarized input beam will in general have a phase difference $\Delta\phi$ between its two components, prior to entering the retarder. Since we are only interested in the geometric phase introduced by the retarder, this input $\Delta\phi$ is defined as a reference condition and we consider only the phase change induced by the retarder.

As we move away from the states whose components split with equal amplitude upon entering the retarder, we find that the geometric phase will no longer be zero. For a 10° linearly polarized input beam, the horizontal and vertical component amplitudes become $\cos(10°)$ and $\sin(10°)$. Ignoring the mean propagation phase, the QWP will displace these components by a phase of $\delta = \pi/2$ relative to each other, so that (6.1) gives $\gamma = 0.7543$ radians (43.22°). For an elliptical polarization with the same amplitude proportion, and ignoring the initial phase difference in the input, we obtain the same geometric phase, $\gamma = 43.22°$. Such states of equal γ form a small circle (a non-great circle) on the Poincaré sphere, also with axis formed by the retarder eigenstates. Thus, the geometric phase obtained on propagating through a linear retarder, for the complete collection of input states, can be visualized as a cylindrically symmetric pattern on the Poincaré sphere, as shown in Fig. 6.4 for the case of three retarder thicknesses — $\lambda/8$, $\lambda/4$, and $\lambda/2$ — all oriented at a fast axis azimuth of 0°.[16]

The advantage of considering the case of a single retarder, beyond starting with a simple example, is that the Jones calculus allows any sequence of polarization components to be concatenated into a single Jones matrix. Thus, any system of retarders can be treated as a single "effective retarder". Consider the case of two QWPs oriented with the same fast axis azimuth angle. Working with (6.1) shows that, unless the polarization state happens to be an eigenstate of the retarder, one cannot simply double the geometric phases of the first QWP to obtain the full geometric phasae obtained by the pair. For non-eigenpolarizations, the polarization state entering the second QWP will be different from that of the first QWP, and so the corresponding γ shift will also be different: $\gamma(\mathrm{QWP}_1) \neq \gamma(\mathrm{QWP}_2)$. At each retarder, the amount of geometric phase introduced depends on

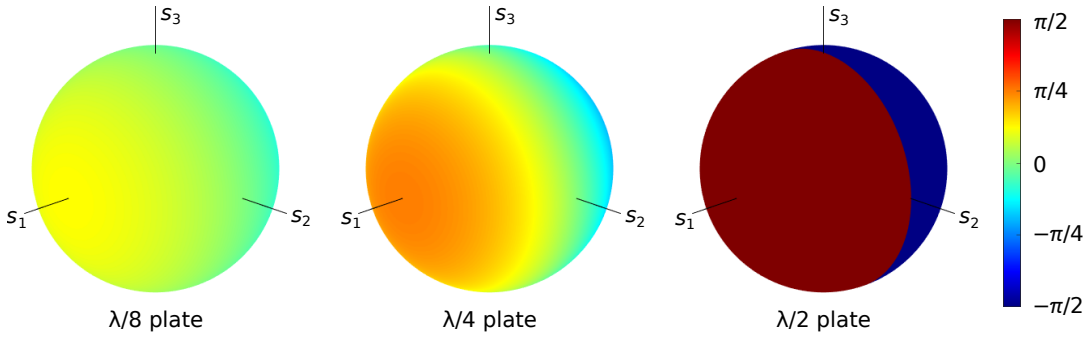

FIGURE 6.4. Geometric phase generated, for all states of polarization, by transmission through a $\lambda/8$, a $\lambda/4$, and a $\lambda/2$ plate, each oriented with fast axis at 0°. All polarization states lying on the $s_1 = 0$ plane have $\gamma = 0$.

the wave's input state of polarization.

If the retarder azimuth is rotated, then the distributions shown in Fig. 6.4 remain the same, except that the pattern rotates such that the axis of symmetry remains aligned to the eigenpolarizations of the waveplate. If we concatenate two QWPs aligned at different azimuth angles, then this system too can be represented as an equivalent retarder, but now with the axis in general aligned along elliptical polarization states. The equivalent retardance δ of the system (of the equivalent retarder representing the QWP pair) can then be inserted into (6.1), together with the appropriate amplitudes A_x and A_y that now represent the amplitudes projected along the eigenstates and not the x and y axes specifically. In fact, any optical system possessing a homogeneous Jones matrix[59] can be treated in exactly this way — insert the equivalent retardance δ and the amplitude projections A_1 and A_2 into (6.1) to determine the geometric phase.

Next we consider the case of concatenating three waveplates (a QHQ) where the QWPs have their fast axis angles at 45°, and the HWP azimuth θ is allowed to rotate freely. The eigenstates of this setup, if we concatenate the three retarders into one equivalent system retarder, are the 0° and 90° linear polarization states. Thus, the QHQ is itself a linear retarder, with equivalent retardance $\delta = 4\theta$.[17] Thus, for incident light oriented along either of the two eigenpolarizations, the output state will have the same state of polarization (SOP) as the input, but acquires a geometric phase equal to $\gamma = 2\theta$. For any other polarization states, the QHQ system transforms the input into a different output state, whose geometric phase can also be calculated by (6.1).

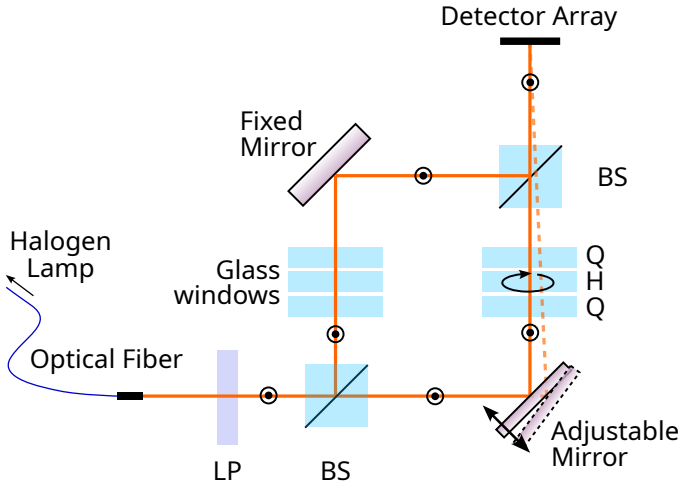

FIGURE 6.5. Mach-Zehnder interferometer layout for adjusting geometric phase (by rotating the half-wave plate inside the QHQ) or for adjusting propagation phase by moving the adjustable mirror. (A photo of the system is shown in Fig. 6.10 in the Appendix.)

6.3 Mach-Zehnder measurements of geometric and propagation phase

While Sec. 6.1 reviewed previous work to explore the boundedness problem using white light interferometry, we developed a similar setup using modern equipment in order to further explore the problem experimentally. At the same time, these experiments provide what we feel to be the clearest yet demonstration of the differences between geometric and propagation phase. In the discussion below, we introduce our experimental setup and show interferogram results for modifying geometric phase or propagation phase. Using these results provides a good basis for clarifying, and explaining, the boundedness problem.

We constructed a white-light Mach-Zehnder interferometer (Fig. 6.5) where the light source is an optical fiber transilluminated by a halogen lamp. The beam passes through a linear polarizer oriented at 90°, which is then split by a non-polarizing beamsplitter. In the measurement arm of the interferometer, we insert a QHQ. In order to equalize the optical path in the two arms, we insert glass plates of appropriate thickness to match the phase. In this case, we used glass plates with a sum total thickness of 12.71 mm. (It is necessary to get the appropriate glass thickness correct to within 0.01 mm.) The two beams interfere at the detector plane at slightly different angles in order to produce a set of transverse fringes. Although the refractive index dispersion of the compensation glass in the reference arm does not exactly match that of the QHQ components in the measurement arm, the match remains close enough to obtain high-contrast fringes and a clear coherence envelope.

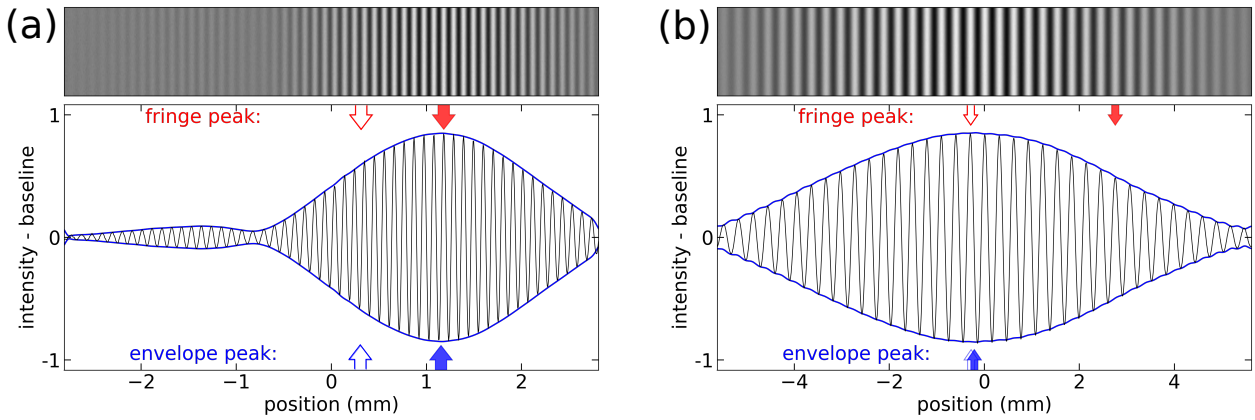

FIGURE 6.6. Mach-Zehnder interferogram for (a) propagation phase change, (b) geometric phase change. The measured interferograms are shown above, and the plots underneath are derived from summing down the columns of the interferogram. The solid arrows indicate the positions of the reference fringe peak (red) and the coherence envelope center (blue) after (1) the adjustable mirror has been moved its full amount, or (a) the QHQ half-wave plate has been rotated about 5 turns. The outline arrows indicate the same features, but show the starting locations, before the mirror has been adjusted. We can see that whereas the interference fringe has moved a lot, the coherence envelope has not.

This Mach-Zender setup has the disadvantages of difficult alignment, and susceptibility to vibrations and changes in airflow. The compensation of optical path in the two arms is also a laborious process. However, it has the advantage that one can easily demonstrate both geometric and propagation phase changes by either rotating the HWP inside the QHQ, or adjusting the position of the fold mirror. Because we do not use a motorized rotational stage for the HWP, or a piezoelectric drive motor for the fold mirror, all of our adjustments are performed manually, while collecting video data at 20 Hz for the geometric phase data, 55 Hz for the propagation phase data. The short duration of the resulting experiment minimizes slow sources of drift in the interferogram, but manual rotation does not allow a constant rotation speed, and it becomes difficult to avoid backlash in the rotational stage. To compensate for this, we can use an algorithm (adapted from Ref. 63) to estimate the mean interferogram bias at each pixel, and also to estimate the phase shift at each frame. Subtracting the bias pixel-by-pixel removes most of the effect produced by blemishes such as dust and optical nonuniformities. Next, we normalize the amplitudes so that the largest amplitudes throughout the datasets are set to $+1$ and -1. Finally, we use the estimated phases to re-organize the video frames to correct for the large amount of jitter caused by the manual adjustment. The result is a high-contrast, smoother video.

In the first (propagation phase) experiment, we collect an interferogram video while displacing

the fold mirror in the measurement arm. In order to make very small displacements, we simply push on the back of the mirror mount using a stiff paper card (standard micrometer stages produce adjustments that are too large for our measurement). While this produces interferogram with a lot of jitter, we once again use the same processing techniques to remove the bias, normalize, and reorder the frames to produce a smoother video. The measurement results for adjusting the propagation phase are shown in Fig. 6.6(a), where the filled red arrow indicates the current position of the fringe, and the hollow red arrow its starting position. Similarly, the filled blue arrow indicates the final position of the coherence envelope center, and the hollow blue arrow its starting position. The interferogram at the top of the figure shows the final frame of the measurement sequence (after the image has been processed to remove bias and to normalize the peak fringe amplitude).

Next we perform the geometric phase experiment by rotating the HWP inside the QHQ by five full turns. Thus, we expect that if we track an individual fringe over the course of the experiment, we should see it move by ten full fringe periods. This is exactly the result shown in Fig. 6.6(b), where once again the filled arrows indicate the final position, and the hollow arrows the starting positions, of the tracked fringe and the coherence envelope. Where as the propagation phase sequence witnesses an approximately equal amount of shift in both the coherence envelope and the tracked fringe — about eight full fringe periods — in the geometric phase case, the coherence envelope does not move while the tracked fringe shifts by about ten full fringe periods.

6.4 Michelson spectral fringe measurements of geometric and propagation phase

In order to visualize the spectral dependence of the fringes directly, we replaced the camera detector array with a line-imaging spectrometer (see Appendix 6.7 for details). Since the Fig. 6.5's Mach-Zehnder setup proved to be problematic for this, we switched to a Michelson interferometer layout, as shown in Fig. 6.7. With this new setup, light passes through the optical elements twice, doubling the amount of geometric phase generated, but unfortunately also doubling the sensitivity for optical path compensation. Doubling of the optical dispersion caused problems for using glass windows to compensate for the optical path length, and so we replaced the glass windows with a second (almost identical) QHQ in the reference arm. This proved sufficient.

The QHQ in the reference arm has its fixed fast axis orientations at 0°for every element, effec-

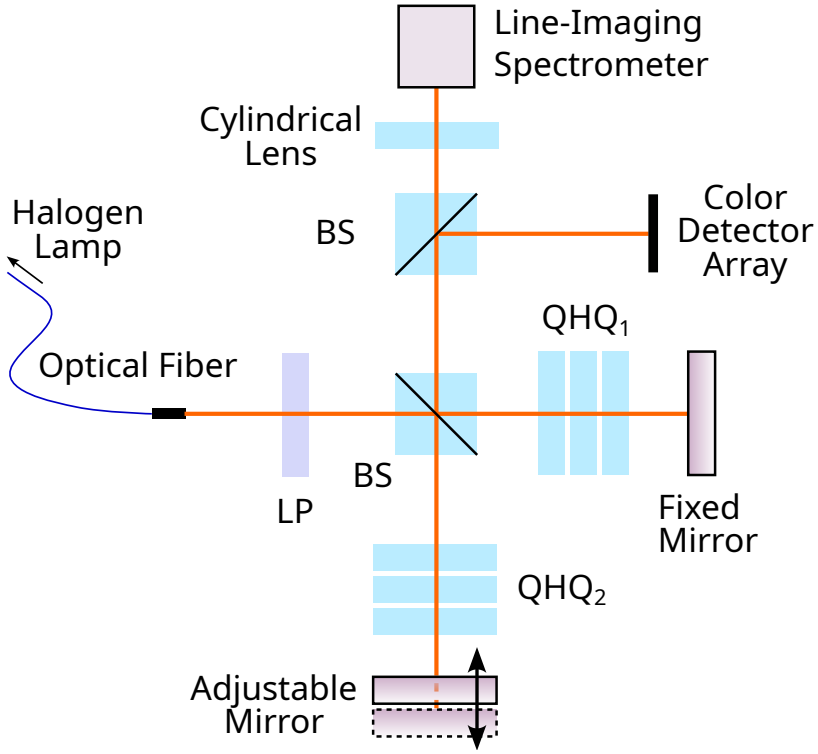

FIGURE 6.7. Michelson interferometer layout for adjusting geometric phase (by rotating the half-wave plate inside QHQ_2) or for adjusting propagation phase by moving the adjustable mirror. (A photo of the system is shown in Fig. 6.11 in the Appendix.)

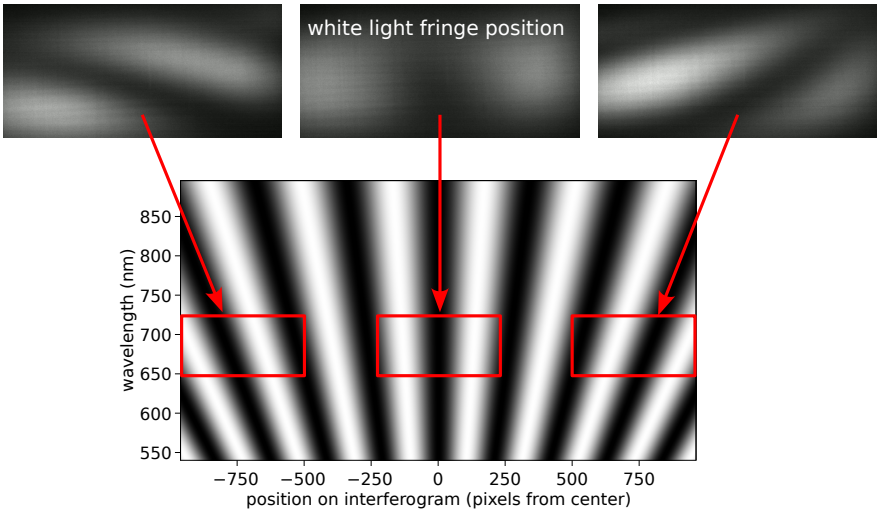

FIGURE 6.8. Three selected frames extracted from the the Michelson interferometer propagation phase measurement sequence captured by the line-imaging spectrometer. As the fold mirror is moved from its original position (the right frame), we see the fringe pattern become upright (at the white light fringe location, the center frame) and then begin tilting in the opposite direction (the left frame). For the geometric phase measurement sequence, the frames look much like that at the far right, with the fringes shifting across the field of view but always returning to their original pattern. The image region corresponding to wavelengths below 550 nm has been cropped for lack of sufficient signal, and to improve the camera frame rate.

tively making it a full-wave plate. The measurement arm QHQ has QWPs fixed at 0°fast axes, and a rotatable HWP at fast axis angle θ. Using this setup, we performed the same approach as with the Mach-Zehnder interferometer to measure the propagation and geometric phase: pushing on the measurement arm mirror, or manually rotating the HWP in the QHQ. Three example frames from these results are shown in Fig. 6.8. Since the fringes in this experiment were quite wide, the figure shows the overall pattern in a larger view (generated by simulation) in order to provide context for the individual frames.

The propagation phase measurement sequence is discussed next. At the beginning, the fringes are tilted to the right, but as we push the measurement arm mirror the shift in phase causes the fringes to rotate, becoming straight (the white light fringe position), and then finally tilting over to the left. The change in fringe tilt reflects the wavelength dependence of propagation phase: shorter wavelengths experience a larger number of modulations for a given amount of mirror motion.

For the geometric phase sequence, we rotate the HWP and observe that the fringe pattern of colors changes cyclically, returning to the same configuration with each half-turn of the HWP. Through the full video sequence, we rotate the HWP through five full turns. As with the Mach-

Zehnder experiments, one can interpret the continuous motion of the fringes from frame to frame as indicating an unbounded phase. Or, if one prefers, one can interpret the cyclical behavior as indicating bounded phase.

6.5 Clarifying the boundedness problem for geometric phase

With our experimental results in hand, we return to consider the boundedness problem, and the puzzling statements made by Bhandari, Hariharan, and their co-workers. Fig. 6.6(b), shows clearly that, for geometric phase changes, the interferogram is periodic. And yet it also shows that if we track the position of any individual fringe in the pattern, its motion is continuous and moves in one direction as we rotate (and will return towards its original location if we start to rotate the HWP in the opposite direction.) This behavior is characteristic of any long-coherence-length source, but in this case we are seeing the behavior with white light.

Translating a fold mirror in the Mach-Zehnder system shifts all wavelengths by an equal amount in path length Δz, but this necessarily means that the shifts if measured in phase units will be different for all wavelengths: $\Delta\phi = \Delta z\, 2\pi/\lambda$. For geometric phase, on the other hand, the shift is approximately equal for all wavelengths if measured in phase $\Delta\phi$, but wavelength-dependent if measured in units of path length Δz.

As Bhandari, Hariharan, Schmitzer, Love, *et al.* have each alluded, one can therefore see the geometric phase as being bounded (the pattern produced is periodic) or as unbounded (the fringe pattern can be shifted as far as one likes) as one desires. However, the shift does not modify the coherence envelope of the light. One cannot, for example, use geometric phase to correct for the path length difference in the two arms of our Mach-Zehnder interferometer: the measurement arm having a QHQ and the reference arm having only air. Without the compensating glass in the reference arm, the peak of the coherence envelope would be far away, and no amount of rotating the HWP inside the QHQ is going to move it any closer. Clearly, if we mix geometric phase and propagation in the same measurement, then a practical perspective would be to say that the geometric phase shift is bounded between $-\pi$ and $+\pi$.

However, there are counter-arguments to this perspective. Consider the trajectory of polarization passing through a thick retarder having many waves of retardance. An input linear polarization at 0°, passing through a thick retarder with fast axis at 45°, will cycle through polarization states

H-R-V-L-H and so on in sequence many times as it propagates through the crystal. Likewise, the geometric phase of the wave will pass through a set of values, with each input state experiencing a different γ value. Now, if we analyze the spectral dependence of γ, then we find that short wavelengths will experience more of these cycles than the long wavelengths will. If we plot γ as a function of wavelength, we obtain a "wrapped phase" according to (6.1). However, the discontinuities of wrapped phase are clearly unphysical — going from one nearby wavelength to another should produce a smooth change in behavior, and not a discontinuous jump. In this case, we make use of the more general form of (6.1) and assume that the value of n is not zero but rather an integer that allows us to restore continuity in the spectrum. This, then, is a practical argument for considering an unbounded phase.

Finally, for systems using a long-coherence light source such as a laser, in which the size of the coherence envelope does not play a significant role in the measurement, the $\pm\pi$ phase bounds place no limitations on the amount of shifting one can do. In this case, a shift of $\phi + 2n\pi$ becomes equivalent to a shift of ϕ, and the continuous tunability allows one to use geometric phase for flexible phase control and active stabilization of an interferometer.[64–66]

In the end, the choice between γ being bounded and unbounded is effectively a choice of convention, much like the choice between the Fresnel or Verdet conventions used in the analysis of optical thin films.[67] The convention may modify the equations we use, and how we visualize the physical behavior of the system, but is really just a choice of what perspective we want to use. The choice does not affect the result, and we need only be careful about what our choice of convention means.

6.6 A choice of convention

In the optics and physics literature, a number of researchers have argued for an unbounded geometric phase.[2–4,62,64,66,68–70] Although we have found no authors arguing explicitly for a bounded phase, the geometric phase literature often contains figures in which the geometric phase is shown as being wrapped.[7,14,43,71] One can consider these figures as implicit arguments for bounded phase. Of course, one can also fall back to the reasonable theoretical argument that any rotation of a linear retarder by 180° should return a system to its original state. By considering both sides of this argument, we feel that the boundedness argument is best considered as a choice of convention.

Perhaps the clearest way to demonstrate the fact that the geometric phase can be considered

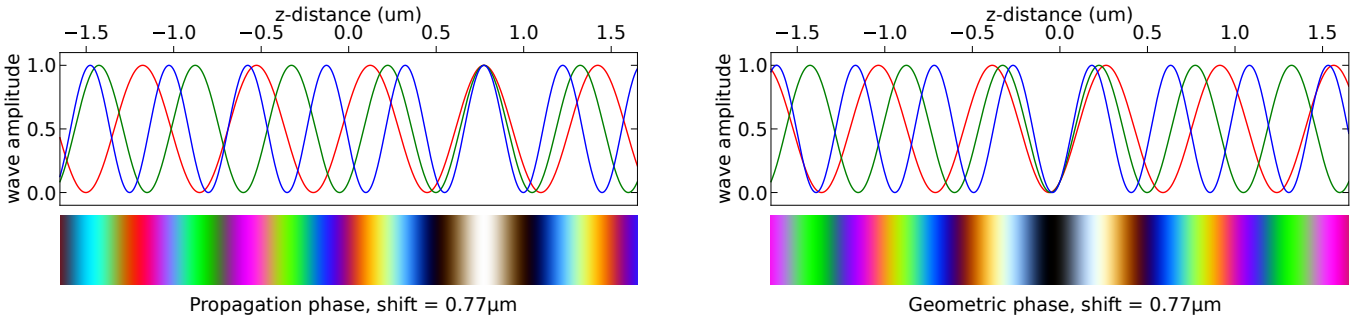

FIGURE 6.9. A simulation of (Upper plots) three laser wavelengths (λ = 450 nm, 550 nm, and 650 nm) combining to create (Lower plots) a simulated white-light interferogram. (Left side) We can see that for a propagation phase shift, the white light fringe moves with the amount of shift (0.77 μm in this case.) (Right side) For a geometric phase shift, the white light fringe stays in place, and merely changes in intensity.

as either bounded or unbounded as one chooses, is the simulation shown in Fig. 6.9. The video shows three waves of different wavelengths that start in the same position. Equal phase shifts are applied to these three waves, but these shifts are either equal in phase (i.e., geometric) or equal in distance (i.e., propagation). As the amount of shift increases, one finds the location of the "white light fringe" stationary for the geometric shifts, but moving for propagation shifts. We hope that this simple simulation, too, can become a useful pedagogical tool.

Although geometric phase is widely described as being achromatic, as indeed our discussion above has also done, it is also important to remember that in practical instruments it is only approximately so. If one constructs a QHQ system as we have done for our experiments, the QWPs and HWP will have wavelength-dependent retardances. Even "achromatic" waveplates are only approximately so. The result is that the paths traced on the Poincaré sphere — or, using the wave description, the evolution of polarization state as the wave passes through the system — will vary with wavelength. We can once again use (6.1) to calculate the geometric phase produced, but note that all of the parameters in the equation (γ, n, δ, A_x, and A_y) will have wavelength dependence.

What makes achromaticity easier to achieve with geometric phase is that, for low-order retarders such as QWPs and HWPs, in contrast to high-order retarders (such as, say, 10λ plates), the sequence of states experienced by the evolving wave are similar even for different wavelengths. To the extent that different wavelengths experience the same evolution of polarization states, they will experience the same geometric phase. If the intermediate SOPs vary strongly with wavelength, then there will be a significant phase dispersion with wavelength. Since we have used achromatic QWPs (although

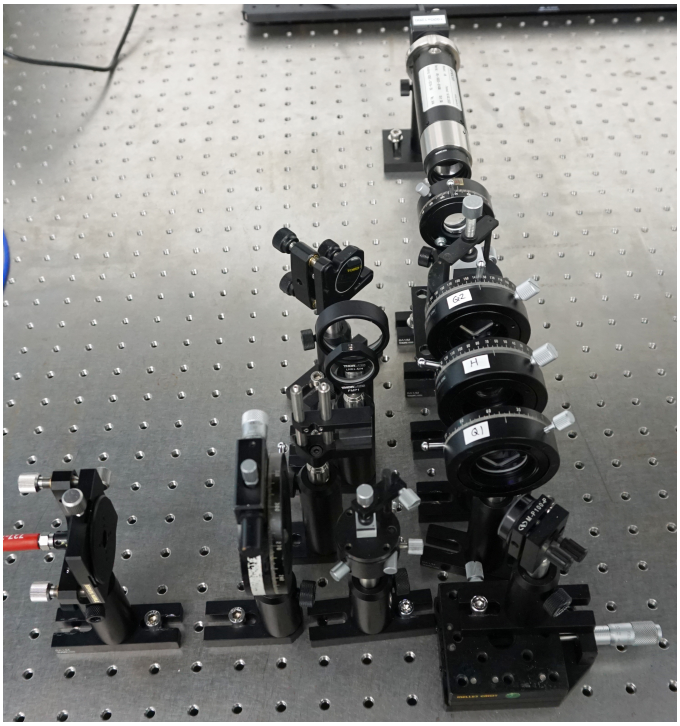

FIGURE 6.10. Photo of the Mach-Zehnder interferometer. In the photo, the detector array (top of photo) used in the experiments has been replaced by a line-imaging spectrometer, but in the experiments the FLIR camera detector array was used.

not achromatic HWPs) in our experiments, the chromatic dependence of γ is largely suppressed, but the polarization state is changing somewhat during the rotations of the HWP.

6.7 Appendix. Experiment details

Figure 6.10 shows a photo of the Mach-Zehnder interferometer experiment setup. The camera used here is a FLIR BFLY-U3-23S6, with a 1920×1200 array of 5.86 µm pixels, capable of operating at 41 Hz. Collecting the full video dataset requires a data flow rate beyond what the computer memory and hard drive can achieve, and so we crop the image to just the central 100 pixels region of the detector array, with $2\times$ binning along the vertical dimension, so that the final dataset has pixel dimensions of 1920×50. Therefore, the final frame rate for video collection is 55 Hz.

The line-imaging spectrometer used in the Michelson interferometer spectral fringe experiments is a Specim Imspector V8, having a 380–800 nm spectral range, and a 50 µm slit width. The detector array attached to the back end of the Specim disperser is the same as used in the Mach-Zehnder experiments — a FLIR BFLY-U3-23S6. We place a cylindrical lens placed in front of the

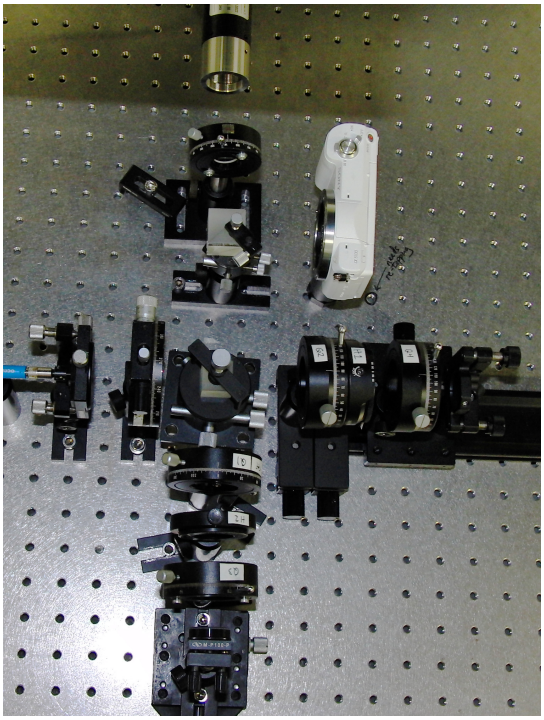

FIGURE 6.11. Photo of the Michelson interferometer. The entrance slit of the line-imaging spectrometer appears at the top.

spectrometer slit (with its long axis oriented horizontally) in order to maximize the light throughput. Once again to reduce memory load, we crop the data to a 884×598 region of the detector array, with $2\times$ binning, reducing the image dimensions to 442×399.

Conclusions

Previous to this thesis the geometric phase was a very difficult to understand concept primarily because of the abstract formalism around it and the lack of a phyisical model. This thesis has shown that the geometric phase can be derived entirely from a concept as fundamental as the superposition of waves. This makes it possible to visualize phase relationships from the waves themselves and understand geometric phase as a shift caused by the relative displacement and amplitudes of the components of a wave. Among the contributions of the wave description of geometric phase it provides a model for calculating γ continuously as a wave propagates through a polarization element, feature that was not available or well understood in the literature. The thesis also presents how an interferogram can have an associated geometric phase and how it can be compared to the wave geometric phase.

Pancharatnam's work shows that if one knows the intensities of two input beams as well as the intensity of their interference, one can infer the phase difference δ between the two input beams from these simple intensity measurements alone. Using the model developed in this thesis, one can reread Pancharatnam's original derivation and see how he considers that the phase of the intensity expression depends on a single cosine. Thus, Pancharatnam is also implicitly considering the case of two cosine waves summing together in the right hand side of his expression for intensity. When he solves for the phase he has a cosine in one side of the equation and a relationship of intensities in the other. With his mathematical prowess he related the intensities to a solid angle in the Poincaré sphere. This is how this thesis arrived to equivalent results through very different conceptual approaches.

Many features of geometric phase that were explored previous to having a physical model are not well understood. One of these is the possible values that geometric phase can have and whether it can be accumulated indefinitely. This thesis shows that the boundedness argument is best considered as a choice of convention. It also reports on experiments that show the differences between the propagation phase and the geometric phase and explains how accumulating retardance modifies the geometric phase, a feature that is not clear with the solid angle approach.

References

[1] Rajendra Bhandari and Joseph Samuel. Observation of topological phase by use of a laser interferometer. *Physical Review Letters*, 60(13):1211, 1988.

[2] Rajendra Bhandari. Observation of non-integrable geometric phase on the poincaré sphere. *Physics Letters A*, 133(1-2):1–3, 1988.

[3] P Hariharan and Maitreyee Roy. A geometric-phase interferometer. 1992.

[4] P Hariharan, Kieran G Larkin, and Maitreyee Roy. The geometric phase: interferometric observations with white light. 1994.

[5] E HECHT 2nd. Optics 2nd ed.[sl], 1990.

[6] R. A. Chipman, W. T. Lam, and G. Young. *Polarized Light and Optical Systems*. CRC Press, 2018.

[7] Chandroth Pannian Jisha, Stefan Nolte, and Alessandro Alberucci. Geometric phase in optics: from wavefront manipulation to waveguiding. *Laser & Photonics Reviews*, 15(10):2100003, 2021.

[8] Shivaramakrishnan Pancharatnam. Generalized theory of interference and its applications, part I. Coherent pencils. *Proceedings of the Indian Academy of Sciences—Section A*, 44(6):398–417, 1956.

[9] M. V. Berry. The adiabatic phase and Pancharatnam's phase for polarized light. *J. Mod. Opt.*, 34:1401–1407, 1987.

[10] M. V. Berry. Quantal phase factors accompanying adiabatic changes. *Proc. Roy. Soc. London A*, 392:45–54, 1984.

[11] Yakir Aharonov and David Bohm. Significance of electromagnetic potentials in the quantum theory. *Physical Review*, 115(3):485, 1959.

[12] Akira Tomita and Raymond Y Chiao. Observation of berry's topological phase by use of an optical fiber. *Physical review letters*, 57(8):937, 1986.

[13] Raymond Y Chiao and Yong-Shi Wu. Manifestations of berry's topological phase for the photon. *Physical review letters*, 57(8):933, 1986.

[14] Thomas Van Dijk, Hugo F Schouten, Wim Ubachs, and Taco D Visser. The pancharatnam-berry phase for non-cyclic polarization changes. *Optics Express*, 18(10):10796–10804, 2010.

[15] Julio C Gutiérrez-Vega. Pancharatnam–berry phase of optical systems. *Optics letters*, 36(7):1143–1145, 2011.

[16] Dorilian Lopez-Mago, Arturo Canales-Benavides, Raul I Hernandez-Aranda, and Julio C Gutiérrez-Vega. Geometric phase morphology of jones matrices. *Optics letters*, 42(14):2667–2670, 2017.

[17] Luis Garza-Soto, Alejandra De-Luna-Pamanes, Israel Melendez-Montoya, Natalia Sanchez-Soria, Diana Gonzalez-Hernandez, and Dorilian Lopez-Mago. Geometric-phase polarimetry. *Journal of Optics*, 22(12):125606, 2020.

[18] Johannes Courtial. Wave plates and the Pancharatnam phase. *Opt. Comm.*, 171:179–183, 1999.

[19] PK Aravind. A simple proof of pancharatnam's theorem. *Optics communications*, 94(4):191–196, 1992.

[20] Zhifan Zhou, Yair Margalit, Samuel Moukouri, Yigal Meir, and Ron Folman. An experimental test of the geodesic rule proposition for the noncyclic geometric phase. *Science advances*, 6(9):eaay8345, 2020.

[21] Piotr Kurzynowski, Władysław A. Woźniak, and Małgorzata Szarycz. Geometric phase: two triangles on the Poincaré sphere. *J. Opt. Soc. Am. A*, 28:475–482, 2011.

[22] José Luis Martínez-Fuentes, Jorge Albero, and Ignacio Moreno. Analysis of optical polarization modulation systems through the pancharatnam connection. *Optics Communications*, 285(4):393–401, 2012.

[23] Julio C Gutiérrez-Vega and Gumaro Rendon. Pancharatnam–berry phase algorithm to calculate the area of arbitrary polygons on the poincaré sphere. *JOSA A*, 37(6):925–929, 2020.

[24] Oriol Arteaga and Hana Bendada. Geometrical phase optical components: measuring geometric phase without interferometry. *Crystals*, 10(10):880, 2020.

[25] Michael David Spivak. A comprehensive introduction to differential geometry. *(No Title)*, 1999.

[26] Rajendra Bhandari. Evolution of light beams in polarization and direction. *Physica B: Condensed Matter*, 175(1-3):111–122, 1991.

[27] Enrique J Galvez. Applications of geometric phase in optics. *Recent Research Developments in Optics*, 2:165–182, 2002.

[28] Enrique J Galvez and Chris D Holmes. Geometric phase of optical rotators. *JOSA A*, 16(8):1981–1985, 1999.

[29] C Cisowski, JB Götte, and S Franke-Arnold. Colloquium: Geometric phases of light: Insights from fiber bundle theory. *Reviews of Modern Physics*, 94(3):031001, 2022.

[30] Eric W Weisstein. *CRC concise encyclopedia of mathematics*. CRC press, 2002.

[31] Rimvydas Aleksiejunas and Vladislovas Ivaska. Geometric phases for scalar wave superpositions. *Physics Letters A*, 235(1):1–6, 1997.

[32] Pacharatnam's connection is named a connection because it can be considered a rule for transporting a vector along a curved space—the surface of the poincaré sphere. in differential geometry, these rules are named connections [spivnak,hannonen].

[33] Surya P Tewari, VS Ashoka, and M Sree Ramana. A four-arm sagnac interferometric switch. *Optics communications*, 120(5-6):235–238, 1995.

[34] Qu Li, Lifu Gong, Yihe Gao, and Yingli Chen. Experimental observation of the nonlinearity of the pancharatnam phase with a michelson interferometer. *Optics communications*, 169(1-6):17–22, 1999.

[35] Bernhard Hils, Wolfgang Dultz, and Werner Martienssen. Nonlinearity of pancharatnam's geometric phase in polarizing interferometers. *Physical Review E*, 60(2):2322, 1999.

[36] H Schmitzer, S Klein, and W Dultz. Nonlinearity of pancharatnam's topological phase. *Physical review letters*, 71(10):1530, 1993.

[37] Antti Hannonen, Kimmo Saastamoinen, Lasse-Petteri Leppänen, Matias Koivurova, Andriy Shevchenko, Ari T Friberg, and Tero Setälä. Geometric phase in beating of light waves. *New Journal of Physics*, 21(8):083030, 2019.

[38] S. Ramaseshan. The Poincare sphere and the Pancharatnam phase — some historical remarks. *Current Science*, 59, 1990.

[39] Rajaram Nityananda, Kausalya Ramaseshan, NV Madhusudana, and GW Series. S pancharatnam (1934–1969): three phases. *Resonance*, 18(4):301–305, 2013.

[40] P. K. Aravind. A simple proof of pancharatnam's theorem. *Opt. Comm.*, 094:191–196, 1992.

[41] S. C. Tiwari. Geometric phase in optics: quantal or classical? *J. Mod. Opt.*, 39:1097–1104, 1992.

[42] M. Roy, P. Svahn, L. Cherel, and C. J. R. Sheppard. Geometric phase-shifting for low-coherence interference microscopy. *Optics and Lasers in Engineering*, 37:631–641, 2002.

[43] José Lages, Remo Giust, and Jean-Marie Vigoureux. Geometric phase and Pancharatnam phase induced by light wave polarization. *Physica E*, 59:6–14, 2014.

[44] Eliahu Cohen, Hugo Larocque, Frédéric Bouchard, Farshad Nejadsattari, Yuval Gefen, and Ebrahim Karimi. Geometric phase from Aharonov-Bohm to Pancharatnam-Berry and beyond. *Nature Rev. Phys.*, 1:437–449, 2019.

[45] Oriol Arteaga. Fresnel-Arago fifth law of interference: the first description of a geometric phase in optics. *J. Mod. Opt.*, 68:350–357, 2021.

[46] Barry Simon. Holonomy, the quantum adiabatic theorem, and Berry's phase. *Phys. Rev. Lett.*, 51:2167–2170, 1983.

[47] S. Ramaseshan and Rajaram Nityananda. The interference of polarized light as an early example of Berry's phase. *Current Science*, 55:1225–1226, 1986.

[48] Michael Berry. Pancharatnam, virtuoso of the Poincaré sphere: an appreciation. *Current Science*, 67:220–223, 1994.

[49] Michael Berry. Geometric phase memories. *Nature Physics*, 6:148–150, 2010.

[50] S. I. Vinitskiĭ, Vladimir Leonardovich Derbov, Vladimir M. Dubovik, B. L. Markovski, and Yu. P. Stepanovskiĭ. Topological phases in quantum mechanics and polarization optics. *Soviet Physics Uspekhi*, 33(6):403, 1990.

[51] J. D. Jackson. Examples of the zeroth theorem of the history of science. *Am. J. Phys.*, 76:704–719, 2008.

[52] E. Collett. *Field Guide to Polarization*. Field Guide Series. Society of Photo Optical, 2005.

[53] Rajaram Nityananda. Pancharatnam's route to the geometric phase. *Current Science*, 67:238–244, 1994.

[54] Isaac Todhunter. *Spherical Trigonometry: For the Use of Colleges and Schools, with Numerous Examples*. Macmillan, 1863.

[55] Julio C. Gutiérrez-Vega. Pancharatnam-Berry phase of optical systems. *Opt. Lett.*, 36:1143–1145, 2011.

[56] Luis Garza-Soto, Nathan Hagen, Dorilian Lopez-Mago, and Yukitoshi Otani. Wave description of geometric phase. *JOSA A*, 40(2):388–396, 2023.

[57] Luis Garza-Soto, Nathan Hagen, and Dorilian Lopez-Mago. Deciphering pancharatnam's discovery of geometric phase: retrospective. *JOSA A*, 40(5):925–931, 2023.

[58] Alexander V Tavrov, Yoko Miyamoto, Tsutomu Kawabata, Mitsuo Takeda, and Vladimir A Andreev. Generalized algorithm for the unified analysis and simultaneous evaluation of geometrical spin-redirection phase and pancharatnam phase in a complex interferometric system. *JOSA A*, 17(1):154–161, 2000.

[59] Julio C Gutiérrez-Vega. Optical phase of inhomogeneous jones matrices: retardance and orthotransmission states. *Optics Letters*, 45(7):1639–1642, 2020.

[60] Rajendra Bhandari. Polarization of light and topological phases. *Physics Reports*, 281(1):1–64, 1997.

[61] H Schmitzer, S Klein, and W Dultz. An experimental test of the path dependency of pancharatnam's geometric phase. *Journal of Modern Optics*, 45(5):1039–1047, 1998.

[62] Gordon D Love. The unbounded nature of geometrical and dynamical phases in polarization optics. *Optics communications*, 131(4-6):236–240, 1996.

[63] Zhaoyang Wang and Bongtae Han. Advanced iterative algorithm for phase extraction of randomly phase-shifted interferograms. *Optics letters*, 29(14):1671–1673, 2004.

[64] Mario Martinelli and Paolo Vavassori. A geometric (pancharatnam) phase approach to the polarization and phase control in the coherent optics circuits. *Optics communications*, 80(2):166–176, 1990.

[65] MU Wehner, MH Ulm, and M Wegener. Scanning interferometer stabilized by use of pancharatnam's phase. *Optics letters*, 22(19):1455–1457, 1997.

[66] P Hariharan. The geometric phase. *Progress in Optics*, 48:149–201, 2005.

[67] RT Holm. Convention confusions. In *Handbook of Optical Constants of Solids*, pages 21–55. Elsevier, 1997.

[68] Y Ben-Aryeh. Berry and pancharatnam topological phases of atomic and optical systems. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(4):R1, 2004.

[69] Rajendra Bhandari. The nonmodular topological phase and phase singularities. *Physics Letters A*, 375(41):3562–3569, 2011.

[70] Michael J Escuti, Jihwan Kim, and Michael W Kudenov. Controlling light with geometric-phase holograms. *Optics and Photonics News*, 27(2):22–29, 2016.

[71] Aleksi Leinonen, Antti Hannonen, Henri Partanen, Janne Heikkinen, Tero Setälä, Ari T Friberg, and Tommi K Hakala. Noncyclic continuous pancharatnam—berry phase in dual-beam interference. *Communications Physics*, 6(1):132, 2023.