Draft version February 2, 2026
Typeset using LaTeX **twocolumn** style in AASTeX7.0.1

An Introduction to Single-Antenna Radio Astronomical Polarimetry

Willem van Straten [1]

[1] *Manly Astrophysics, 15/41-42 East Esplanade, Manly, NSW 2095, Australia*

Submitted to Publications of the Astronomical Society of the Pacific

ABSTRACT

This tutorial reviews the mathematical foundations of single-antenna radio polarimetry with the aim of fostering a conceptual understanding of the relationships between a physical description of signal propagation (gain, delay, reflection, down-conversion, etc.), the corresponding transformations of the electric field vector, and the equivalent operations on the Stokes parameters. The adopted framework is based on the work of Britton (2000) and Hamaker (2000) and applied to analyze the signal path described by Hamaker et al. (1996) with additional corrections for phase convention and reflection. Some objective criteria for selecting a model of the instrumental response are introduced and discussed, along with some practical guidelines that facilitate polarimetric calibration. Further relevant background material and lengthier mathematical proofs are included in the appendix, which introduces the vector, matrix, and tensor notation and concepts of linear algebra used in this work. The appendix also reviews some of the basics of analog and digital signal processing that are relevant to radio astronomy, and discusses some numerical instabilities that arise when modeling observations.

Keywords: Polarimetry (1278)

Contents

1. INTRODUCTION

The polarization of electromagnetic radiation reveals otherwise unattainable details about a wide variety of astrophysical phenomena, from high-energy cosmic ray air showers (Aab et al. 2014; Schellart et al. 2014) to solar coronal mass ejections (Kansabanik et al. 2024) and the relativistic precession of neutron stars (Desvignes et al. 2019). On larger scales, radio polarization has been used to study the plasma in the vicinity of a supermassive black hole (Event Horizon Telescope Collaboration et al. 2021), and the energy density of gravitational waves generated during inflation (Polnarev 1985; Planck Collaboration et al. 2020).

As a signal from an astrophysical source propagates to its point of reception, its state of polarization is altered, which can be used to study the physical properties of the media through which the signal has traveled. For example, observations of the polarized emission from radio pulsars are used to study the Earth's ionosphere (Porayko et al. 2023) and map the large-scale structure of the Galactic magnetic field (Beck et al. 1996; Han et al. 2018). By revealing the strength and orientation of magnetic fields, polarization observations are fundamental to understanding their role in star formation, galaxy evolution, and high-energy phenomena.

Measuring polarization requires purpose-built instrumentation known as a *polarimeter*, and the experimental activities related to measuring polarization are known as *polarimetry*. A polarimeter typically distorts the polarization state of the signal in a manner that is unintended and cannot be deduced from theory. This unknown component of the instrumental response to an electromagnetic signal must be determined experimentally and calibrated before the polarization intrinsic to the astrophysical source can be interpreted.

Methods of polarimetric calibration are based on a mathematical model that describes the propagation of a polarized signal along the path between the source and the point of detection. Accordingly, the primary aim of this paper is to foster a conceptual understanding of the mathematical foundations of polarimetry. The insights gained through this approach facilitate the development of practical guidelines for use when designing an experiment or analyzing observations.

This review focuses on polarimetric observations of point sources made using a single antenna, such as a single dish or a phased array formed by the phase-coherent addition of signals from multiple elements of an interferometric array. For a comprehensive introduction to both the theory and practice of radio polarimetry using interferometric arrays, including rigorous treatment of arbitrary brightness distributions and direction-dependent

effects, the pioneering series by Smirnov (2011a,b,c,d) and the lessons learned by Lenc et al. (2017) are highly recommended reading. The formalism employed by Smirnov and the framework presented in this tutorial are both heavily influenced by Hamaker (2000).

Before embarking on this approach, Section 2 of this paper reviews the geometric description of polarization introduced by Stokes (1852), beginning with an ideal, monochromatic source of electromagnetic radiation. In Section 3, the Stokes parameters are related to the second-order statistics of the electric field as represented by the coherency matrix. This connection is further explored in Section 4, where linear transformations of the electric field are related to Euclidean rotations and Lorentz boosts of the Stokes parameters. In Section 5, these transformations are applied to a decomposition of the signal path that extends the Hamaker et al. (1996) model to include the effects of reflection and down-conversion. Inverting such transformations to recover the original polarization state is discussed in Section 6, which includes a first-order calibration solution based on the ideal feed assumptions. Section 7 introduces some criteria for selecting a more complete model of the instrumental response, and Section 8 concludes with some pointers to further reading.

Throughout this article, mathematical expressions are greatly simplified by using complex-valued tensors, such as vectors and matrices, and the various operations that are performed on them. These are described in the appendix with an introduction to the fundamentals of multivariate linear algebra. The appendix also reviews the basics of signal processing, such as the analytic signal and down-conversion. It concludes with some discussion of numerical stability when modeling polarization observations.

2. MONOCHROMATIC LIGHT

At radio frequencies, it is both possible to directly sample the electric field, and sufficient to employ a classical description of electromagnetic radiation (Maxwell 1865; Born & Wolf 1980). At higher frequencies in the electromagnetic spectrum, polarimeters incorporate devices that count photons, such as charge-coupled devices and photomultiplier tubes. For a comprehensive introduction to the quantum mechanical treatment of the polarization of light, please see the excellent review by Goldberg et al. (2021).

In both classical and quantum approaches to the topic, it is useful to develop some intuition for the geometry of polarization by starting with the nonphysical, idealized case of a purely monochromatic, plane-propagating, transverse electromagnetic wave. For such a wave, there exist two linearly-independent solutions to Maxwell's equations, representing two oppositely polarized waves. Therefore, to fully describe the vector state of any observed signal, radio receiver systems are designed with a pair of receptors that are ideally sensitive to orthogonal senses of polarization. Typically, the receptors are either circularly polarized (e.g., left and right circularly polarized) or linearly polarized (e.g., horizontally and vertically polarized).

The following discussion is based on a Cartesian coordinate system in which the electromagnetic wave propagates in the positive z direction and the electric field is completely described by two orthogonal linearly-polarized components in the x-y plane, which is perpendicular to the direction of wave propagation,

$$\boldsymbol{\epsilon}(t) = \begin{pmatrix} \epsilon_x(t) \\ \epsilon_y(t) \end{pmatrix} = \begin{pmatrix} a_x \cos(2\pi\nu t + \phi_x) \\ a_y \cos(2\pi\nu t + \phi_y) \end{pmatrix}. \quad (1)$$

Here, ϵ_x and ϵ_y are the real-valued projections of $\boldsymbol{\epsilon}$ onto the x and y axes. In the special case of monochromatic light, ν is the constant frequency of the wave, a_x and a_y are the constant amplitudes of the orthogonal wave components, and ϕ_x and ϕ_y are the phases of these components at $t = 0$.

On each cycle of the wave, the electric field vector traces an ellipse in the x–y plane, as shown in Figure 1. This figure depicts two angles that describe the geometry of this ellipse: the position angle ψ and the ellipticity angle χ. Stokes (1852) first demonstrated the utility of expressing the components of the electric field vector in terms of these angles, starting with an intermediate reference frame that is rotated about the z-axis by the position angle ψ, such that

$$x' = x\cos\psi + y\sin\psi \quad (2)$$

$$y' = -x\sin\psi + y\cos\psi. \quad (3)$$

In this reference frame, which is depicted with dashed lines and open arrow heads in Figure 1, the electric field is expressed as in equation (2) of Stokes (1852)

$$\boldsymbol{\epsilon}'(t) = \begin{pmatrix} \epsilon'_x(t) \\ \epsilon'_y(t) \end{pmatrix} = \begin{pmatrix} r\cos\chi\sin(2\pi\nu t + \phi) \\ r\sin\chi\cos(2\pi\nu t + \phi) \end{pmatrix}, \quad (4)$$

where r defines the size of the ellipse and ϕ defines the position of the electric field vector at $t = 0$.

When $\chi = 0$, $\epsilon'_y = 0$ and the electric field vector oscillates along a line defined by the x' axis, which is oriented with respect to the x axis by ψ. When $\chi > 0$, the electric field travels in a clockwise direction about the ellipse, as seen by an observer looking toward the source, which is defined as a left-hand polarized wave (IEEE 1983).

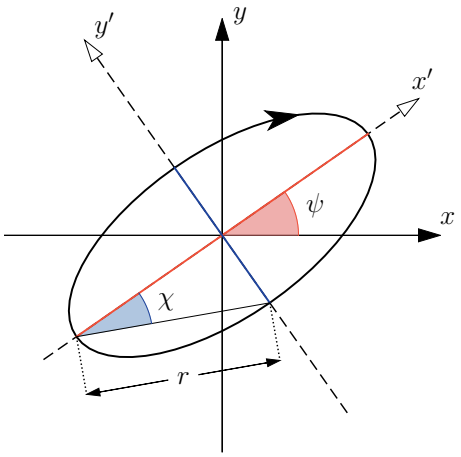

Figure 1. A left-handed ellipse is traced by the electric field vector on each cycle of a monochromatic wave travelling in the positive z direction, as observed when looking toward the source. The position angle, $-\pi/2 \leq \psi \leq \pi/2$, describes the orientation of the semi-major axis of the ellipse with respect to the x axis; it is positive when the x' and y' axes are rotated in a counter-clockwise direction with respect to the x and y axes. The tangent of the ellipticity angle, $-\pi/4 \leq \chi \leq \pi/4$, is equal to the ratio between the semi-minor axis and the semi-major axis; it is positive for a left-hand polarized wave and negative for a right-hand polarized wave. The size of the ellipse is defined by r, which is equal to the length of the hypotenuse of the right triangle that contains χ.

When $\chi < 0$, the counter-clockwise traverse of the electric field vector is defined as right-hand polarization. In the special cases of $\chi = \pm\pi/4$, the amplitudes of ϵ'_x and ϵ'_y are equal and the electric field traces a circle in the x–y plane. In general, polarization state depends on both the amplitudes of the electric field components and the relative phase between them.

To move beyond the nonphysical case of perfectly monochromatic waves, for which amplitude and phase are constant as a function of time, it is necessary to characterize random fluctuations of both amplitude and phase using statistical averages. This is facilitated by adopting the complex-valued analytic representation of a signal, which directly encodes its instantaneous amplitude and phase (see Appendix A.1). After replacing the real-valued $\epsilon_x(t)$ and $\epsilon_y(t)$ with their associated analytic signals, $e_x(t)$ and $e_y(t)$, the monochromatic wave described by Equations (1) through (4) is given by the real part of its analytic representation,

$$\boldsymbol{e}(t) = \begin{pmatrix} e_x(t) \\ e_y(t) \end{pmatrix} = \boldsymbol{e}_0 \exp(i[2\pi\nu t + \phi - \pi/2]). \quad (5)$$

Here, the polarization state of the wave is completely described by the complex-valued two-dimensional Jones vector (Eqn. D20),

$$\boldsymbol{e}_0 = r \begin{pmatrix} \cos\psi\cos\chi - i\sin\psi\sin\chi \\ \sin\psi\cos\chi + i\cos\psi\sin\chi \end{pmatrix}. \quad (6)$$

An interactive notebook[1] demonstrates the relation between $\boldsymbol{\epsilon}(t)$ and $\boldsymbol{e}(t)$ by depicting the ellipses inscribed by these functions as r, ψ and χ are varied.

3. PARTIAL POLARIZATION

No physical signal is ever strictly monochromatic, and astrophysical signals are typically an incoherent superposition of waves from a large number of emitting sources. The polarization of such signals cannot be described as in the previous section. Furthermore, over most of the spectrum, the electric field fluctuates too rapidly to be directly sampled, and it is generally not possible to determine the geometry of the polarization ellipse directly from measurements of the electric field vector. Instead, the polarization state must be inferred by other means.

For example, the position angles of X-ray photons are determined from the trajectories of ejected photoelectrons. The polarization of optical light is typically estimated by measuring pairs of intensities after the radiation is decomposed into oppositely polarized streams. (See Appendix E for more detail.) At radio frequencies, the electric field is directly sampled; however, the astrophysical signal of interest is typically buried in noise and it is necessary to integrate over time and frequency to increase the signal-to-noise ratio. Such integration can be performed only after squaring the electric field, after which the information about the instantaneous phase of the signal is lost.

3.1. *Stokes Parameters*

Although polarization state is not measured directly from the electric field, its interpretation remains founded on the geometry of the polarization ellipse owing to the analysis by Stokes (1852). By considering the conditions under which two independent sources of radiation could be considered equivalent, Stokes (1852) arrived at four measurable quantities that completely describe the state of polarization. Following refinements by Perrin (1942), the four Stokes parameters are given by

$$\begin{aligned} S_0 &= I \\ S_1 &= Ip\cos 2\chi\cos 2\psi \\ S_2 &= Ip\cos 2\chi\sin 2\psi \\ S_3 &= Ip\sin 2\chi \end{aligned} \quad (7)$$

[1] https://github.com/straten/epsic/tree/master/notebooks

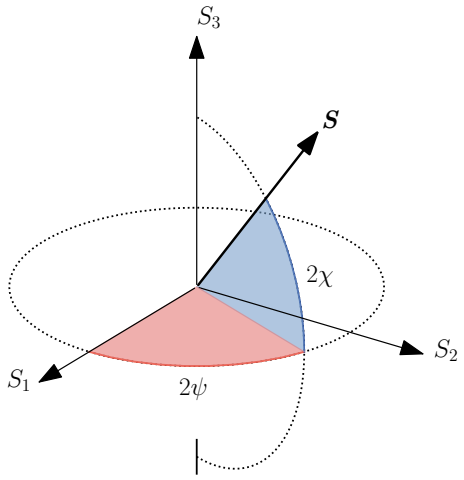

Figure 2. The spherical coordinates of the polarization vector, $\underline{\boldsymbol{S}} = (S_1, S_2, S_3)^T = (Q, U, V)^T$, include the vector length $|\underline{\boldsymbol{S}}| = Ip$, longitude 2ψ, and latitude 2χ. Points in the S_1–S_2 plane ($S_3 = 0$) represent linearly polarized states, points above this plane ($S_3 > 0$) represent left-hand elliptically polarized states and the positive S_3 pole represents left-hand circular polarization.

where I is the intensity of the light, p is the degree of polarization (see Section 3.2), and χ and ψ are the ellipticity and orientation angles that define the polarization ellipse.

The four-dimensional vector of Stokes parameters,

$$\underline{\underline{S}} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}, \tag{8}$$

is separable into scalar and vector components, $[S_0, \underline{\boldsymbol{S}}]$, where $\underline{\boldsymbol{S}} = (S_1, S_2, S_3)^T$ is the polarization vector with length $|\underline{\boldsymbol{S}}| = Ip$ and direction defined by 2ψ and 2χ, as depicted in Figure 2. Measurement of $\underline{\boldsymbol{S}}$ yields the orientation and ellipticity angles that define the geometry of the polarization ellipse,

$$\psi = \frac{1}{2} \tan^{-1} \frac{S_2}{S_1} \tag{9}$$

$$\chi = \frac{1}{2} \sin^{-1} \frac{S_3}{|\underline{\boldsymbol{S}}|}. \tag{10}$$

3.2. *Degree of Polarization*

The degree of polarization

$$p = \frac{|\underline{\boldsymbol{S}}|}{S_0} \tag{11}$$

is a measure of the fraction of the total intensity that is polarized. When $p = 0$, the light is unpolarized, also defined as *common* light (Stokes 1852); when $p = 1$, the light is 100% polarized, or *purely* polarized; for partial polarization, $0 < p < 1$.

As shown by Equation (C17) and the discussion that follows it, unpolarized or partially polarized radiation is equivalent to an incoherent superposition of two orthogonal purely polarized states. An unpolarized state can be decomposed into any pair of orthogonal purely polarized states of equal amplitude (Stokes 1852). In contrast, a 100% polarized wave can be described using a single Jones vector, $\boldsymbol{e}_0$, such that

$$\boldsymbol{e}(t) = \boldsymbol{e}_0 z(t), \tag{12}$$

where $z(t)$ is a complex-valued function that describes stochastic fluctuations of amplitude and phase. It has a mean of zero and variance of unity; that is, $\langle |z(t)|^2 \rangle = 1$, where angular brackets denote an average over time. Referring to Equations (5) and (6), monochromatic light is a special case of Equation (12) in which $z(t)$ is a pure tone with constant frequency; therefore, strictly monochromatic light is 100% polarized.

3.3. *Poincaré Sphere*

When only the polarization state is of interest, the polarization vector can be normalized by the total intensity, yielding $\underline{\acute{\boldsymbol{S}}} = \underline{\boldsymbol{S}}/S_0$, such that $|\underline{\acute{\boldsymbol{S}}}| = p$. When $\underline{\acute{\boldsymbol{S}}}$ is plotted as in Figure 2, every possible polarization state corresponds to a point on or inside a sphere of unity-radius known as the Poincaré sphere. In this geometric representation, purely polarized states lie on the surface of the sphere, and the unpolarized state is at its center (the origin). Partially polarized states are represented by points within the sphere, and their distance from the center is directly proportional to their degree of polarization.

3.4. *Polarization Ellipse*

In this section, the relationship between the measurable Stokes parameters and the inferred geometry of the polarization ellipse described by Equation (7) is explored using the special case of purely polarized monochromatic light described by Equation (1). The connection between the *relative* phases[2] and amplitudes of the components of the electric field vector, the geometry of the ellipse traced by that vector, and the corresponding Stokes parameters are demonstrated using six special cases of purely polarized radiation: Stokes ± Q, ± U, and ± V. These cases are depicted in Figure 3.

[2] The Stokes parameters are independent of the absolute phase, ϕ, that appears in Equation (4). That is, the measured polarization state does not depend on the choice of time origin.

Stokes $\pm Q$: If $a_y = 0$ in Equation (1), then the radiation is 100% linearly polarized with the electric field vector oscillating along the x axis. In this case, $\chi = 0$, $\psi = 0$, Stokes $Q = S_1 = S_0$ is positive, and Stokes $S_2 = S_3 = 0$.

If $a_x = 0$, then the radiation is 100% linearly polarized with electric field vector oscillating along the y axis. In this case, $\chi = 0$, $\psi = \pm\pi/2$, Stokes $Q = S_1 = -S_0$ is negative, and Stokes $S_2 = S_3 = 0$.

Stokes $\pm U$: If $a_x = a_y$ and the relative phase difference, $\Delta\phi = \phi_x - \phi_y = 0$, then $\epsilon_x(t) = \epsilon_y(t)$ and the radiation is 100% linearly polarized with the electric field vector oscillating along the line defined by $y = x$ (i.e., offset by 45° with respect to the x axis). In this case, $\chi = 0$, $\psi = \pi/4$, Stokes $U = S_2 = S_0$ is positive, and Stokes $S_1 = S_3 = 0$.

If $a_x = a_y$, and $\Delta\phi = \pm\pi$, then $\epsilon_x(t) = -\epsilon_y(t)$ and the radiation is 100% linearly polarized with the electric field vector oscillating along the line defined by $y = -x$ (i.e., offset by $-45°$ with respect to the x axis). In this case, $\chi = 0$, $\psi = -\pi/4$, Stokes $U = S_2 = -S_0$ is negative, and Stokes $S_1 = S_3 = 0$.

Stokes $\pm V$: If $a_x = a_y$ and $\Delta\phi = -\pi/2$, then the phase of ϵ_y leads that of ϵ_x by 90° and the radiation is 100% circularly polarized with the electric field vector tracing a clockwise circle in the x−y plane, as seen by an observer looking toward the source; this is defined as a left-hand circularly polarized (LCP) wave. In this case, $\chi = \pi/4$, ψ is undefined, Stokes $V = S_3 = S_0$ is positive, and Stokes $S_1 = S_2 = 0$.

If $a_x = a_y$ and $\Delta\phi = \pi/2$, then the phase of ϵ_x leads that of ϵ_y by 90° and the radiation is 100% circularly polarized with the electric field vector tracing a counterclockwise circle in the x−y plane, as seen by an observer looking toward the source; this is defined as a right-hand circularly polarized (RCP) wave. In this case, $\chi = -\pi/4$, ψ is undefined, Stokes $V = S_3 = -S_0$ is negative, and Stokes $S_1 = S_2 = 0$.

3.5. *Adopted Standard*

Apart from the refinements described in Appendix F, the conventions and definitions adopted in this article are consistent with both Stokes (1852) and the Institute of Electrical and Electronics Engineers (IEEE 1983). In the IEEE standard, the right-hand rule is applied with the thumb pointing in the direction of wave propagation; it defines both the angle measured from a reference axis to the major axis of the polarization ellipse, and the direction in which the electric field rotates for right-hand elliptically polarized states.

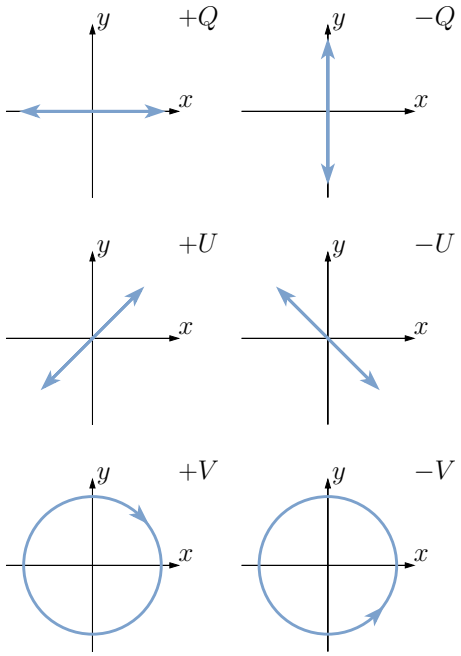

Figure 3. The polarization ellipses inscribed by the electric field vector for the six special cases of monochromatic light described in Section 3.4.

The IEEE standard also defines the Poincaré sphere such that left-hand elliptically polarized states occupy the upper hemisphere, where Stokes V is positive in a right-handed basis defined by Stokes Q, U and V. As more fully discussed in Hamaker & Bregman (1996) and van Straten et al. (2010), the International Astronomical Union (IAU 1974) defines Stokes $V = -S_3$, which is negative for left-handed circular polarization and opposite to the convention adopted in this article.

3.6. *Coherency Matrix*

The polarization state of electromagnetic radiation can also be described by the second-order statistics of $\boldsymbol{e}(t)$, as represented by the complex-valued 2×2 coherency matrix (Born & Wolf 1980)

$$\boldsymbol{\rho} \equiv \langle \boldsymbol{e}(t) \otimes \boldsymbol{e}^\dagger(t) \rangle = \begin{pmatrix} \langle e_0 e_0^* \rangle & \langle e_0 e_1^* \rangle \\ \langle e_1 e_0^* \rangle & \langle e_1 e_1^* \rangle \end{pmatrix}. \quad (13)$$

Here, the angular brackets denote an average over time, $\otimes$ is the tensor product (Eqn. B11), and $\boldsymbol{e}^\dagger(t)$ is the Hermitian transpose of $\boldsymbol{e}(t)$ (Eqn. B3). For brevity, explicit dependence on time has been dropped in the rightmost expression. Note that each component of $\boldsymbol{e}(t)$ is multiplied by the complex conjugate of either itself or the other component, a process known as *square law detection*. Consequently, as for the Stokes parameters, the components of the coherency matrix are independent of the absolute phase of $\boldsymbol{e}(t)$.

The coherency matrix is self-adjoint, or Hermitian (i.e., $\boldsymbol{\rho} = \boldsymbol{\rho}^\dagger$), and can be written as a linear combination of four Hermitian basis matrices (e.g., Fano 1957),

$$\boldsymbol{\rho} = \frac{1}{2}\sum_{\mu=0}^{3} S_\mu \boldsymbol{\sigma}_\mu, \tag{14}$$

where S_μ are the four real-valued Stokes parameters and the basis matrices consist of the 2×2 identity matrix and the Pauli matrices,

$$\begin{aligned} \boldsymbol{\sigma}_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \boldsymbol{\sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ \boldsymbol{\sigma}_2 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \boldsymbol{\sigma}_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \end{aligned} \tag{15}$$

For brevity in the remainder of this paper, the summation symbol is omitted from equations and Einstein notation is used to imply a sum over repeated indeces. (See Appendix B for more detail.)

Conversely, the Stokes parameters can be represented as projections of the coherency matrix onto the basis matrices using the tensor double contraction (Eqn. B13),

$$S_\mu = \boldsymbol{\sigma}_\mu : \boldsymbol{\rho}. \tag{16}$$

Proof 3.1.

$$\begin{aligned} \boldsymbol{\sigma}_\mu : \boldsymbol{\rho} &= \frac{1}{2} S_\kappa \, \boldsymbol{\sigma}_\mu : \boldsymbol{\sigma}_\kappa && \text{Eqn. (14)} \\ &= S_\kappa \delta_{\mu\kappa} && \text{Eqn. (C4)} \\ &= S_\mu \end{aligned}$$

Substitution of Equation (13) into Equation (16) leads to the following expressions for the Stokes parameters.

$$S_0 = \langle |e_0(t)|^2 \rangle + \langle |e_1(t)|^2 \rangle \tag{17}$$
$$S_1 = \langle |e_0(t)|^2 \rangle - \langle |e_1(t)|^2 \rangle \tag{18}$$
$$S_2 = 2\mathrm{Re}\left[\langle e_0^*(t) e_1(t) \rangle\right] \tag{19}$$
$$S_3 = 2\mathrm{Im}\left[\langle e_0^*(t) e_1(t) \rangle\right] \tag{20}$$

Example: Consider $\mu = 3$,

$$\begin{aligned} S_3 &= \boldsymbol{\sigma}_3 : \boldsymbol{\rho} && \text{Eqn. (16)} \\ &= \boldsymbol{\sigma}_3 : \langle \boldsymbol{e} \otimes \boldsymbol{e}^\dagger \rangle && \text{Eqn. (13)} \\ &= \mathrm{Tr}\left[\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \langle e_0 e_0^* \rangle & \langle e_0 e_1^* \rangle \\ \langle e_1 e_0^* \rangle & \langle e_1 e_1^* \rangle \end{pmatrix} \right] && \text{Eqn. (B15)} \\ &= -i\langle e_1 e_0^* \rangle + i \langle e_0 e_1^* \rangle \\ &= 2\,\mathrm{Im}[\langle e_0^* \, e_1 \rangle]. \end{aligned}$$

3.7. *Invariant Interval*

As for the Lorentz four-vector in special relativity, the square of the invariant interval of the Stokes parameters is defined by (Barakat 1963; Britton 2000)

$$|\underline{\underline{\boldsymbol{S}}}|^2 = S_0^2 - |\underline{\boldsymbol{S}}|^2. \tag{21}$$

This interval remains invariant (up to scalar multiples) under linear transformations of the electric field; therefore, $|\underline{\underline{\boldsymbol{S}}}|$ is linearly proportional to the scalar amplification of the source. This property proves useful during numerical analysis and modeling; for example, the invariant can be used to normalize the Stokes parameters and accurately compensate for random scintillation-induced fluctuations of the flux density of the source (as in van Straten 2004). In high-precision pulsar timing experiments, use of the invariant interval yields arrival time estimates with significantly reduced instrumental artifacts (e.g., van Straten et al. 2001).

Through its linear relation to the determinant of the coherency matrix via Equation (C6),

$$|\boldsymbol{\rho}| = |S_\mu \boldsymbol{\sigma}_\mu / 2| = (S_0^2 - |\underline{\boldsymbol{S}}|^2)/4 = |\underline{\underline{\boldsymbol{S}}}|^2/4, \tag{22}$$

the invariant also helps to define the concept of a valid polarization state. As a direct consequence of the Cauchy-Schwarz inequality for random variables, a valid polarization state must satisfy

$$|\boldsymbol{\rho}| \geq 0 \quad \text{and} \quad |\underline{\underline{\boldsymbol{S}}}| \geq 0. \tag{23}$$

Proof 3.2. The Cauchy-Schwarz inequality,

$$|\langle e_0 e_1^* \rangle|^2 \leq \langle |e_0|^2 \rangle \langle |e_1|^2 \rangle,$$

and Equation (13) yield

$$|\boldsymbol{\rho}| = \langle |e_0|^2 \rangle \langle |e_1|^2 \rangle - |\langle e_0 e_1^* \rangle|^2 \geq 0$$

Any polarization state that fails to satisfy Equation (23) may be called *invalid*, *nonphysical*, and/or *over-polarized*. Although nonphysical, over-polarized states can arise during numerical analysis. For example, they can be produced through transformation by an impure Mueller matrix (e.g. Section 7.4). They can also arise when all four Stokes parameters are allowed to vary independently during modeling.

For fully polarized light, $S_0 = |\underline{\boldsymbol{S}}|$ and $|\underline{\underline{\boldsymbol{S}}}| = |\boldsymbol{\rho}| = 0$. The invariant interval and determinant are also zero for the *instantaneous* Stokes parameters,

$$\tilde{s}_\mu \equiv \boldsymbol{\sigma}_\mu : \tilde{\boldsymbol{\rho}} \quad \text{where} \quad \tilde{\boldsymbol{\rho}} \equiv \boldsymbol{e} \otimes \boldsymbol{e}^\dagger, \tag{24}$$

because $\tilde{\boldsymbol{\rho}}$ is computed from a single instance of the electric field and is therefore singular; i.e.,

$$|\tilde{\boldsymbol{\rho}}| = |\boldsymbol{e} \otimes \boldsymbol{e}^\dagger| = e_0 e_0^* \, e_1 e_1^* - e_0 e_1^* \, e_1 e_0^* = 0. \tag{25}$$

3.8. *Orthogonal Polarizations*

Stokes (1852) defined oppositely polarized waves by considering the decomposition of a partially polarized signal into two purely polarized components,

$$\boldsymbol{e}_A(t) = \boldsymbol{e}_a z_a(t) \quad \text{and} \quad \boldsymbol{e}_B(t) = \boldsymbol{e}_b z_b(t), \tag{26}$$

where $\boldsymbol{e}_a$ and $\boldsymbol{e}_b$ are constant Jones vectors that represent the polarizations of the two components, and $z_i(t) = \boldsymbol{e}_i^\dagger \cdot \boldsymbol{e}(t)$ ($i \in \{a, b\}$) are the projections of $\boldsymbol{e}(t)$ onto these vectors (Eqns. [B3] and [B4]). He then considered the superposition of the two components after applying a relative delay τ between them,

$$\boldsymbol{e}'(t) = \boldsymbol{e}_A(t) + \boldsymbol{e}_B(t - \tau), \tag{27}$$

and defined $\boldsymbol{e}_A$ and $\boldsymbol{e}_B$ as oppositely polarized if and only if the total intensity of $\boldsymbol{e}'(t)$ is independent of τ. The total intensity is given by (see Eqn. 17),

$$I = \langle \boldsymbol{e}^\dagger(t) \cdot \boldsymbol{e}(t) \rangle = \langle |e_0(t)|^2 + |e_1(t)|^2 \rangle; \tag{28}$$

therefore, the total intensity of the superposition

$$\begin{aligned} I' &= \langle [\boldsymbol{e}_A(t) + \boldsymbol{e}_B(t-\tau)]^\dagger \cdot [\boldsymbol{e}_A(t) + \boldsymbol{e}_B(t-\tau)] \rangle \\ &= \boldsymbol{e}_a^\dagger \cdot \boldsymbol{e}_a \langle |z_a(t)|^2 \rangle + \boldsymbol{e}_b^\dagger \cdot \boldsymbol{e}_b \langle |z_b(t-\tau)|^2 \rangle \\ &\qquad + 2\mathrm{Re}\left[\boldsymbol{e}_a^\dagger \cdot \boldsymbol{e}_b \, \langle z_a(t) z_b^*(t-\tau) \rangle \right]. \end{aligned} \tag{29}$$

Assuming that the signal is stationary, at least in the weak sense, $\langle |z_b(t-\tau)|^2 \rangle = \langle |z_b(t)|^2 \rangle$; therefore, I' is independent of τ, and the two components are orthogonal, if and only if $\boldsymbol{e}_a^\dagger \cdot \boldsymbol{e}_b = 0$.

To characterize the Stokes parameters of orthogonally polarized states, consider the singular coherency matrices of the two purely polarized states,

$$\tilde{\boldsymbol{\rho}}_a \equiv \boldsymbol{e}_a \otimes \boldsymbol{e}_a^\dagger \quad \text{and} \quad \tilde{\boldsymbol{\rho}}_b \equiv \boldsymbol{e}_b \otimes \boldsymbol{e}_b^\dagger,$$

and their tensor double contraction,

$$\tilde{\boldsymbol{\rho}}_a : \tilde{\boldsymbol{\rho}}_b = |\boldsymbol{e}_a^\dagger \cdot \boldsymbol{e}_b|^2. \tag{30}$$

Proof 3.3.

$$\begin{aligned} \tilde{\boldsymbol{\rho}}_a : \tilde{\boldsymbol{\rho}}_b &= \mathrm{Tr}\left[\tilde{\boldsymbol{\rho}}_a \, \tilde{\boldsymbol{\rho}}_b\right] && \text{Eqn. (B15)} \\ &= \mathrm{Tr}\left[\boldsymbol{e}_a \boldsymbol{e}_a^\dagger \, \boldsymbol{e}_b \boldsymbol{e}_b^\dagger\right] && \text{Eqn. (B12)} \\ &= \mathrm{Tr}\left[\boldsymbol{e}_a^\dagger \boldsymbol{e}_b \boldsymbol{e}_b^\dagger \, \boldsymbol{e}_a\right] && \text{Eqn. (B15)} \\ &= \mathrm{Tr}\left[(\boldsymbol{e}_a^\dagger \boldsymbol{e}_b)(\boldsymbol{e}_a^\dagger \boldsymbol{e}_b)^*\right] \\ &= |\boldsymbol{e}_a^\dagger \boldsymbol{e}_b|^2 && \mathrm{Tr}\,[z] = z \end{aligned}$$

Therefore, the coherency matrices of two purely polarized sources are orthogonal with respect to the trace inner product (Eqn. B15) when the polarizations are orthogonal. Given the associated Stokes parameters,

$$\tilde{s}_{a,\mu} = \boldsymbol{\sigma}_\mu : \tilde{\boldsymbol{\rho}}_a \quad \text{and} \quad \tilde{s}_{b,\mu} = \boldsymbol{\sigma}_\mu : \tilde{\boldsymbol{\rho}}_b,$$

the tensor double contraction,

$$\tilde{\boldsymbol{\rho}}_a : \tilde{\boldsymbol{\rho}}_b = \frac{1}{2} \left(\tilde{s}_{a,0} \tilde{s}_{b,0} + \underline{\tilde{\boldsymbol{s}}}_a \cdot \underline{\tilde{\boldsymbol{s}}}_b \right). \tag{31}$$

Proof 3.4.

$$\begin{aligned} \tilde{\boldsymbol{\rho}}_a : \tilde{\boldsymbol{\rho}}_b &= (\tilde{s}_{a,\mu} \, \boldsymbol{\sigma}_\mu / 2) : (\tilde{s}_{b,\nu} \, \boldsymbol{\sigma}_\nu / 2) \\ &= \frac{1}{4} \tilde{s}_{a,\mu} \, \tilde{s}_{b,\nu} \, \boldsymbol{\sigma}_\mu : \boldsymbol{\sigma}_\nu \\ &= \frac{1}{2} \tilde{s}_{a,\mu} \, \tilde{s}_{b,\nu} \, \delta_{\mu\nu} && \text{Eqn. (C4)} \\ &= \frac{1}{2} \tilde{s}_{a,\mu} \, \tilde{s}_{b,\mu} \end{aligned}$$

Equating the right-hand sides of Equations (30) and (31), noting that $\tilde{s}_0 = |\underline{\tilde{\boldsymbol{s}}}|$, and rearranging yields

$$\underline{\tilde{\boldsymbol{s}}}_a \cdot \underline{\tilde{\boldsymbol{s}}}_b = 2|\boldsymbol{e}_a^\dagger \cdot \boldsymbol{e}_b|^2 - |\underline{\tilde{\boldsymbol{s}}}_a||\underline{\tilde{\boldsymbol{s}}}_b|. \tag{32}$$

The angle between $\underline{\tilde{\boldsymbol{s}}}_a$ and $\underline{\tilde{\boldsymbol{s}}}_b$ is given by

$$\cos\Theta = \frac{\underline{\tilde{\boldsymbol{s}}}_a \cdot \underline{\tilde{\boldsymbol{s}}}_b}{|\underline{\tilde{\boldsymbol{s}}}_a||\underline{\tilde{\boldsymbol{s}}}_b|} = \frac{2|\boldsymbol{e}_a^\dagger \cdot \boldsymbol{e}_b|^2}{|\underline{\tilde{\boldsymbol{s}}}_a||\underline{\tilde{\boldsymbol{s}}}_b|} - 1. \tag{33}$$

If $\boldsymbol{e}_a^\dagger \cdot \boldsymbol{e}_b = 0$, then $\cos\Theta = -1$ and the angle between $\underline{\tilde{\boldsymbol{s}}}_a$ and $\underline{\tilde{\boldsymbol{s}}}_b$ is 180°. That is, the polarization vectors of orthogonally polarized signals are anti-parallel.

This result is consistent with Stokes (1852), where it is demonstrated that oppositely polarized waves trace ellipses with orthogonal major axes ($|\psi_A - \psi_B| = \pi/2$) and equal and opposite handedness ($\chi_A = -\chi_B$). That is, on the Poincaré sphere, oppositely polarized waves occupy antipodal points.

4. LINEAR TRANSFORMATIONS

Polarimetric studies require modeling and correcting the manner in which the measurement apparatus alter the incident radiation. In radio astronomy, the instrument includes everything from the antenna (including any reflectors involved in focusing and redirecting the radio waves) to the signal processing system used to compute the Stokes parameters. It is also necessary to model and correct Faraday rotation, which occurs in the Earth's ionosphere and the interstellar medium. In this section, the response of a system is described using linear transformations of the electric field and the equivalent linear transformations of the Stokes parameters.

4.1. *Jones Matrices*

In the narrow-band (or quasi-monochromatic) approximation of an electromagnetic wave (see Appendix A.2), the response of a single receptor is defined by a complex-valued row vector, $\boldsymbol{r} = (r_0, r_1)$, such that the voltage induced in the receptor by the incident electric field $\boldsymbol{e}(t)$ is given by the scalar product, $v(t) = \boldsymbol{r} \cdot \boldsymbol{e}(t)$. The intensity of the response, $I = \langle v^2(t) \rangle$, is maximum when the polarization of the incident wave matches that of the receptor, and reduces to zero when the incident wave is orthogonally polarized.

The response of a dual-receptor feed is represented by a 2×2 complex-valued Jones matrix with rows equal to the receptor vectors,

$$\mathbf{J} = \begin{pmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{pmatrix} = \begin{pmatrix} r_{00} & r_{01} \\ r_{10} & r_{11} \end{pmatrix}, \tag{34}$$

such that any linear transformation of the electric field vector may be represented by

$$\boldsymbol{e}'(t) = \mathbf{J}\boldsymbol{e}(t). \tag{35}$$

The receptors in an ideal feed are orthogonal, such the scalar product $\boldsymbol{r}_0 \cdot \boldsymbol{r}_1^\dagger = 0$.

The response of a system that is composed of a series of elements is represented as a product of Jones matrices. Matrix multiplication is not commutative, and in general the order in which operations are performed must correctly reflect the order in which the elements in the signal path are encountered. For example, consider the system response described by equation (11) of Hamaker et al. (1996),

$$\boldsymbol{e}'(t) = \mathbf{J}\boldsymbol{e}(t) = \mathbf{G}\,\mathbf{D}\,\mathbf{C}\,\mathbf{P}\,\mathbf{F}\,\boldsymbol{e}(t). \tag{36}$$

From the incident electric field $\boldsymbol{e}(t)$ on the right to the observed $\boldsymbol{e}'(t)$ on the left, astrophysical signals are subjected to

- Faraday rotation in both the interstellar medium and the ionosphere, **F**;
- the projection between the celestial reference frame and the receptor basis, **P**;
- the nominal antenna and feed configuration, **C**;
- deviations from an ideal feed, **D**; and
- complex receiver gains, **G**.

These transformations are depicted in Figure 4 and defined and discussed in more detail in Section 5. Note that Figure 4 includes an additional phase convention correction $\underline{\underline{\mathbf{\Phi}}}$ that cannot be represented using a Jones matrix and therefore does not appear in Equation (36). The phase convention correction can be represented by a Mueller matrix and is included in Equation (42) of Section 4.3.

4.2. *Congruence Transformations*

Substitution of $\boldsymbol{e}' = \mathbf{J}\boldsymbol{e}$ into Equation (13) yields the coherency matrix of the transformed electric field, which is described by a congruence transformation,

$$\boldsymbol{\rho}' = \mathbf{J}\boldsymbol{\rho}\mathbf{J}^\dagger. \tag{37}$$

Proof 4.1.

$$\begin{aligned}
\boldsymbol{\rho}' &= \langle \boldsymbol{e}' \otimes \boldsymbol{e}'^\dagger \rangle && \text{Eqn. (13)} \\
&= \langle (\mathbf{J}\boldsymbol{e}) \otimes (\mathbf{J}\boldsymbol{e})^\dagger \rangle && \text{Eqn. (35)} \\
&= \langle \mathbf{J}\boldsymbol{e} \otimes \boldsymbol{e}^\dagger \mathbf{J}^\dagger \rangle && (\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger \\
&= \mathbf{J} \langle \boldsymbol{e} \otimes \boldsymbol{e}^\dagger \rangle \mathbf{J}^\dagger && \mathbf{J} \text{ is constant} \\
&= \mathbf{J}\boldsymbol{\rho}\,\mathbf{J}^\dagger && \text{Eqn. (13)}
\end{aligned}$$

As discussed in more detail in Section 4.5, the congruence transformation is a special case of the more general radio interferometer measurement equation for a single point source (Hamaker 2000; Smirnov 2011a). For a single antenna, the coherency matrix is insensitive to the absolute phase of **J**.

Proof 4.2. Consider $\mathbf{J}' = \mathbf{J}e^{i\phi}$.

$$\begin{aligned}
\boldsymbol{\rho}' &= \mathbf{J}'\boldsymbol{\rho}\mathbf{J}'^\dagger && \text{Eqn. (37)} \\
&= \left(\mathbf{J}e^{i\phi}\right) \boldsymbol{\rho} \left(\mathbf{J}e^{i\phi}\right)^\dagger \\
&= e^{i\phi}\mathbf{J}\boldsymbol{\rho}\mathbf{J}^\dagger e^{-i\phi} = \mathbf{J}\boldsymbol{\rho}\mathbf{J}^\dagger
\end{aligned}$$

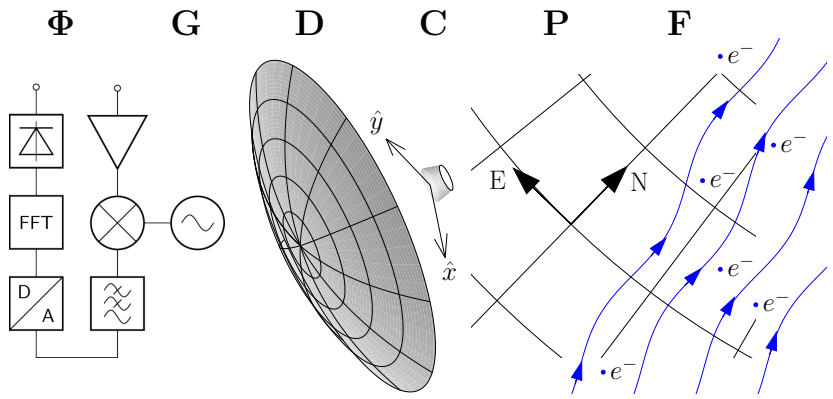

Figure 4. The radio wave signal path, from free space propagation on the right to detection on the left. Also from right to left are Faraday rotation **F** due to free electrons permeated by a magnetic field (represented by blue lines with arrows); the coordinate projection **P** between the x and y axes of the receiver and the North and East basis vectors of the celestial reference frame; the nominal configuration **C** of an ideal antenna and feed; unintended deviations **D** from that ideal; and the gains **G** and phase convention $\underline{\underline{\mathbf{\Phi}}}$ of the down-conversion system and instrumentation used to detect the Stokes parameters.

Noting that $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger$, congruence transformation of the coherency matrix by the Jones matrix defined in Equation (36) has the following "onion" form.

$$\boldsymbol{\rho}' = \mathbf{G}\left(\mathbf{D}\left(\mathbf{C}\left(\mathbf{P}\left(\mathbf{F}\boldsymbol{\rho}\mathbf{F}^\dagger\right)\mathbf{P}^\dagger\right)\mathbf{C}^\dagger\right)\mathbf{D}^\dagger\right)\mathbf{G}^\dagger \tag{38}$$

(The matrices in this equation are defined and discussed in more detail in Section 5.)

Polarimetric calibration consists of transforming the observed coherency matrix by the inverse of the system response, $\boldsymbol{\rho} = \mathbf{J}^{-1}\boldsymbol{\rho}'\mathbf{J}^{-1\dagger}$, yielding the intrinsic polarization,

$$\boldsymbol{\rho} = \cdots\mathbf{C}^{-1}\left(\mathbf{D}^{-1}\left(\mathbf{G}^{-1}\boldsymbol{\rho}\mathbf{G}^{-1\dagger}\right)\mathbf{D}^{-1\dagger}\right)\mathbf{C}^{-1\dagger}\cdots \tag{39}$$

Because $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, the onion is turned inside out and operations are inverted in reverse order, beginning with the last element of the signal path in the innermost congruence transformation.

A congruence transformation converts any valid polarization state to another valid state.

Proof 4.3. If $|\boldsymbol{\rho}| \geq 0$ (Eqn. 23), then

$$\begin{aligned}
|\boldsymbol{\rho}'| &= |\mathbf{J}\,\boldsymbol{\rho}\,\mathbf{J}^\dagger| && \text{Eqn. (37)}\\
&= |\mathbf{J}||\boldsymbol{\rho}||\mathbf{J}^\dagger| && |\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|\\
&= JJ^*|\boldsymbol{\rho}| && J = |\mathbf{J}|\\
&\geq 0 && |\boldsymbol{\rho}| \geq 0 \text{ and } zz^* \geq 0
\end{aligned}$$

No linear transformation of the electric field can alter the degree of polarization of a purely polarized state.

Proof 4.4. If $\boldsymbol{\rho}$ is purely polarized, then

$$\begin{aligned}
|\boldsymbol{\rho}'| &= |\mathbf{J}\,\boldsymbol{\rho}\,\mathbf{J}^\dagger| && \text{Eqn. (37)}\\
&= |\mathbf{J}||\boldsymbol{\rho}||\mathbf{J}^\dagger| && |\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|\\
&= 0 && |\boldsymbol{\rho}| = 0
\end{aligned}$$

Conversely, only a singular transformation ($|\mathbf{J}| = 0$) transforms partial polarization into pure polarization.

4.3. *Mueller Matrices*

Using Equations (14) and (16), a congruence transformation of the coherency matrix can be expressed as an equivalent linear transformation of the associated Stokes parameters by a real-valued 4×4 Mueller matrix $\underline{\underline{\mathbf{M}}}$, as defined by

$$S'_\mu = M^\nu_\mu S_\nu \tag{40}$$

where

$$M^\nu_\mu = \frac{1}{2}\boldsymbol{\sigma}_\mu : \left(\mathbf{J}\boldsymbol{\sigma}_\nu\mathbf{J}^\dagger\right). \tag{41}$$

Proof 4.5.

$$\begin{aligned}
S'_\mu &= \boldsymbol{\sigma}_\mu : \boldsymbol{\rho}' && \text{Eqn. (16)}\\
&= \boldsymbol{\sigma}_\mu : \left(\mathbf{J}\boldsymbol{\rho}\mathbf{J}^\dagger\right) && \text{Eqn. (37)}\\
&= \boldsymbol{\sigma}_\mu : \left(\mathbf{J}\left[S_\nu\,\boldsymbol{\sigma}_\nu/2\right]\mathbf{J}^\dagger\right) && \text{Eqn. (14)}\\
&= \frac{1}{2}\boldsymbol{\sigma}_\mu : \left(\mathbf{J}\boldsymbol{\sigma}_\nu\,\mathbf{J}^\dagger\right)S_\nu && S_\nu \text{ is a scalar}\\
&= M^\nu_\mu S_\nu && \text{Eqn. (41)}
\end{aligned}$$

Although there is a unique Mueller matrix for every Jones matrix, the converse is not true. This is most directly understood by noting that a 4×4 real-valued matrix has 16 degrees of freedom (dof), and a 2×2

complex-valued matrix has only 8 dof. Mueller matrices that do not have an equivalent Jones matrix are known as "impure" or "depolarizing" (e.g., Hamaker et al. 1996; Lu & Chipman 1996, and Appendix C.7). The definition of a pure Mueller matrix is derived in Appendix C.6.

The signal path described by van Straten et al. (2010) extends the Hamaker et al. (1996) model with an impure correction for the complex phase convention, which is described in more detail in Section 5.1. Therefore, the van Straten et al. (2010) calibration model is best summarized using the equivalent transformation of the Stokes parameters by a Mueller matrix $\underline{\underline{\mathbf{M}}}$,

$$\underline{\boldsymbol{S}}' = \underline{\underline{\mathbf{M}}}\,\underline{\boldsymbol{S}} = \underline{\underline{\mathbf{\Phi}}}\,\underline{\underline{\mathbf{G}}}\,\underline{\underline{\mathbf{D}}}\,\underline{\underline{\mathbf{C}}}\,\underline{\underline{\mathbf{P}}}\,\underline{\underline{\mathbf{F}}}\,\underline{\boldsymbol{S}}. \tag{42}$$

From right to left, $\underline{\boldsymbol{S}}$ is the column vector with elements equal to the 4 Stokes parameters intrinsic to the astronomical source; $\underline{\underline{\mathbf{F}}}$ through $\underline{\underline{\mathbf{G}}}$ are the Mueller matrices derived from the Jones matrices defined in Equation (36) and discussed in Section 5; $\underline{\underline{\mathbf{\Phi}}}$ is the phase convention correction (see Section 5.1); and $\underline{\boldsymbol{S}}'$ is the column vector of observed Stokes parameters.

4.4. *Polarization Transfer Tensors*

Linear transformations of polarization state can also be represented as a double contraction (Eqn. B13) with a two-dimensional, rank 4 polarization transfer tensor, $\boldsymbol{\mathcal{U}}$, such that

$$\boldsymbol{\rho}' = \boldsymbol{\mathcal{U}} : \boldsymbol{\rho}. \tag{43}$$

Similar tensors are used in the generalized radio interferometer equation (Smirnov 2011d) and in descriptions of the propagation of radiation through a magnetized plasma (e.g., Kawabata 1964; Zheleznyakov 1968). The tensor formalism also simplifies the analysis of the covariances between the Stokes parameters (van Straten & Tiburzi 2017) and the definition of pure Mueller matrices (Appendix C.6).

To relate a congruence transformation of the coherency matrix to the equivalent double contraction with a polarization transfer tensor, the $\tilde{\otimes}$ operator is introduced to represent a tensor product followed by a transpose over covariant[3] tensor indeces (Cardoso 1991). That is, where $\mathbf{A}$ and $\mathbf{B}$ are matrices (rank 2 tensors) and

$$\{\mathbf{A} \otimes \mathbf{B}\}^{jl}_{ik} \equiv A^j_i B^l_k \tag{44}$$

is their tensor product,

$$\{\mathbf{A} \,\tilde{\otimes}\, \mathbf{B}\}^{jl}_{ik} \equiv A^l_i B^j_k. \tag{45}$$

[3] Covariant tensor indeces are raised in this work; note that van Straten & Tiburzi (2017) incorrectly identified this as a transpose over contravariant tensor indeces.

Upon double contraction with a matrix $\mathbf{C}$, the rank 4 tensors defined by Equations (44) and (45) exhibit the following transformation properties (Proofs B.1 and B.2).

$$(\mathbf{A} \otimes \mathbf{B}) : \mathbf{C} = \mathbf{A}\,(\mathbf{B} : \mathbf{C}) \tag{46}$$

$$(\mathbf{A} \,\tilde{\otimes}\, \mathbf{B}) : \mathbf{C} = \mathbf{ACB} \tag{47}$$

Using Equations (43) and (47), a congruence transformation by a Jones matrix $\mathbf{J}$ can be expressed as a double contraction with

$$\boldsymbol{\mathcal{U}} = \mathbf{J} \,\tilde{\otimes}\, \mathbf{J}^\dagger. \tag{48}$$

Proof 4.6.

$$\begin{aligned}
\boldsymbol{\rho}' &= \boldsymbol{\mathcal{U}} : \boldsymbol{\rho} && \text{Eqn. (43)} \\
&= \left(\mathbf{J} \,\tilde{\otimes}\, \mathbf{J}^\dagger\right) : \boldsymbol{\rho} && \text{Eqn. (48)} \\
&= \mathbf{J}\,\boldsymbol{\rho}\,\mathbf{J}^\dagger. && \text{Eqn. (47)}
\end{aligned}$$

The polarization transfer tensor can also replace the congruence transformation in Equation (41), thereby yielding the equivalent Mueller matrix $\underline{\underline{\mathbf{M}}}$,

$$M^\nu_\mu = \frac{1}{2}\boldsymbol{\sigma}_\mu : \boldsymbol{\mathcal{U}} : \boldsymbol{\sigma}_\nu. \tag{49}$$

The inverse of this mapping is given by

$$\boldsymbol{\mathcal{U}} = \frac{1}{2} M^\nu_\mu \boldsymbol{\sigma}_\mu \otimes \boldsymbol{\sigma}_\nu. \tag{50}$$

Proof 4.7.

$$\begin{aligned}
&\frac{1}{2}\boldsymbol{\sigma}_\mu : \boldsymbol{\mathcal{U}} : \boldsymbol{\sigma}_\nu && \text{Eqn. (49)} \\
&= \frac{1}{2}\boldsymbol{\sigma}_\mu : \left(\frac{1}{2} M^\lambda_\kappa \boldsymbol{\sigma}_\kappa \otimes \boldsymbol{\sigma}_\lambda\right) : \boldsymbol{\sigma}_\nu && \text{Eqn. (50)} \\
&= \frac{1}{4} M^\lambda_\kappa \boldsymbol{\sigma}_\mu : (\boldsymbol{\sigma}_\kappa \otimes \boldsymbol{\sigma}_\lambda) : \boldsymbol{\sigma}_\nu \\
&= \frac{1}{4} M^\lambda_\kappa \,(\boldsymbol{\sigma}_\mu : \boldsymbol{\sigma}_\kappa)\,(\boldsymbol{\sigma}_\lambda : \boldsymbol{\sigma}_\nu) && \text{Eqn. (46)} \\
&= M^\lambda_\kappa \delta_{\mu\kappa}\delta_{\nu\lambda} && \text{Eqn. (C4)} \\
&= M^\nu_\mu
\end{aligned}$$

Equation (49) expresses the components of $\underline{\underline{\mathbf{M}}}$ as the double projections of $\boldsymbol{\mathcal{U}}$ onto the Hermitian basis matrices. Equation (50) represents $\boldsymbol{\mathcal{U}}$ as a linear combination of the 16 basis tensors formed by tensor products of the 4 Hermitian basis matrices.

4.5. *Comparison with Radio Interferometry*

As in Smirnov (2011a), the radio interferometer measurement equation for a single point source models the visibility matrices formed by the outer products of electric field vectors from pairs of elements in an array. That is, if $\boldsymbol{e}_A(t)$ and $\boldsymbol{e}_B(t)$ are the signals from two array elements, the visibility matrix for this pair

$$\boldsymbol{V}_{AB} \equiv \langle \boldsymbol{e}_A(t) \otimes \boldsymbol{e}_B^\dagger(t) \rangle. \tag{51}$$

Furthermore, if $\mathbf{J}_A$ and $\mathbf{J}_B$ are the Jones matrices that describe the reception of the incident electric field $\boldsymbol{e}(t)$ by the two array elements, such that $\boldsymbol{e}_A(t) = \mathbf{J}_A \boldsymbol{e}(t)$ and $\boldsymbol{e}_B(t) = \mathbf{J}_B \boldsymbol{e}(t)$, then the observed visibility matrix,

$$\boldsymbol{V}_{AB} = \mathbf{J}_A \, \boldsymbol{\rho} \, \mathbf{J}_B^\dagger. \tag{52}$$

In the context of interferometry, the coherency matrix $\boldsymbol{\rho}$ is also known as the brightness matrix (Smirnov 2011a). Equation (52) reduces to Equation (37) when observing a single point source with a single antenna. In the tensor formalism of the generalized radio interferometer measurement equation (Smirnov 2011d), Equation (52) is expressed as

$$\boldsymbol{V}_{AB} = \boldsymbol{\mathcal{U}}_{AB} : \boldsymbol{\rho}, \tag{53}$$

where $\boldsymbol{\mathcal{U}}_{AB} = \mathbf{J}_A \,\tilde{\otimes}\, \mathbf{J}_B^\dagger$ is the visibility transfer tensor[4] (cf. Eqn. 48).

An extended source is described by a brightness matrix that varies with direction, $\boldsymbol{\rho}(\hat{\underline{\boldsymbol{z}}})$, such that the visibility matrix is given by the integral

$$\boldsymbol{V}_{AB} = \iint \mathbf{J}_A(\hat{\underline{\boldsymbol{z}}}) \, \boldsymbol{\rho}(\hat{\underline{\boldsymbol{z}}}) \, \mathbf{J}_B^\dagger(\hat{\underline{\boldsymbol{z}}}) \, d\Omega, \tag{54}$$

where $\mathbf{J}_{A,B}(\hat{\underline{\boldsymbol{z}}})$ are the direction-dependent responses of each antenna. For a single-antenna observation of an extended source,

$$\boldsymbol{\rho} = \iint \mathbf{J}(\hat{\underline{\boldsymbol{z}}}) \, \boldsymbol{\rho}(\hat{\underline{\boldsymbol{z}}}) \, \mathbf{J}^\dagger(\hat{\underline{\boldsymbol{z}}}) \, d\Omega. \tag{55}$$

There are some key differences between single-antenna and interferometric measurements of polarization. For a single antenna, the coherency matrix is Hermitian and independent of the absolute phase of $\boldsymbol{e}(t)$. For an interferometer, the visibility matrix is not Hermitian and depends on the relative phase between $\boldsymbol{e}_A(t)$ and $\boldsymbol{e}_B(t)$.

[4] A transpose of covariant tensor indeces also appears in equation (10) of Smirnov (2011d).

4.6. *Polar Decomposition*

Any invertible Jones matrix $\mathbf{J}$ can be decomposed into a unique product known as its left polar decomposition,

$$\mathbf{J} = J \, \mathbf{B}\mathbf{R}, \tag{56}$$

where $J = |\mathbf{J}|^{1/2}$ and $|\mathbf{J}|$ is the determinant of $\mathbf{J}$; $\mathbf{B}$ is Hermitian ($\mathbf{B}^\dagger = \mathbf{B}$); $\mathbf{R}$ is unitary ($\mathbf{R}^\dagger = \mathbf{R}^{-1}$); and both $\mathbf{B}$ and $\mathbf{R}$ are unimodular ($|\mathbf{B}| = |\mathbf{R}| = 1$). The requirement for a left polar decomposition[5] stems from the definition of a Jones matrix as a pair of receptor row vectors (Eqn. 34).

Polar decomposition into Hermitian and unitary matrices provides a practical framework for classifying and conceptualizing distinct polarization-altering effects, as discussed in more detail in Section 4.7. Isolating the determinant of the Jones matrix also has practical benefit. In single-antenna polarimetry, the coherency matrix is insensitive to the phase of J (Proof 4.2); therefore, without any loss of generality, the absolute phase is set to zero and J is replaced by the real-valued absolute gain, $G = |J|$, thereby eliminating a degenerate dof.

The polar decomposition also enables unique determination of the Hermitian component of the instrumental response given only an observation of a source that is known to be completely unpolarized. To demonstrate this, first note that the Hermitian component of any Jones matrix $\mathbf{J}$ can be determined via the Gram matrix of its rows, $\mathbf{J}\mathbf{J}^\dagger$, such that[6]

$$\mathbf{B}^2 = G^{-2} \mathbf{J}\mathbf{J}^\dagger. \tag{57}$$

Proof 4.8. The Gram matrix of the rows of $\mathbf{J}$,

$$\begin{aligned} \mathbf{J}\mathbf{J}^\dagger &= J \, \mathbf{B}\,\mathbf{R} \, (J \, \mathbf{B}\,\mathbf{R})^\dagger \\ &= J \, \mathbf{B}\,\mathbf{R}\,\mathbf{R}^\dagger \mathbf{B}^\dagger J^* \\ &= |J|^2 \mathbf{B}\,\mathbf{R}\,\mathbf{R}^{-1}\mathbf{B} \\ &= G^2 \, \mathbf{B}^2 \end{aligned}$$

For an unpolarized source $\boldsymbol{\rho} = I\boldsymbol{\sigma}_0$; therefore,

$$\boldsymbol{\rho}' = \mathbf{J}\boldsymbol{\rho}\,\mathbf{J}^\dagger = I\mathbf{J}\mathbf{J}^\dagger = IG^2\mathbf{B}^2. \tag{58}$$

[5] A right polar decomposition, $\mathbf{J}' = J' \, \mathbf{R}' \, \mathbf{B}'$, would be used if receptors were column vectors, such that $\boldsymbol{e}'(t) = \mathbf{J}'^\dagger \boldsymbol{e}(t)$ and $\boldsymbol{\rho}' = \mathbf{J}'^\dagger \boldsymbol{\rho} \, \mathbf{J}'$. Note that $\mathbf{J}' = \mathbf{J}^\dagger = J^* \, \mathbf{R}^\dagger \mathbf{B}^\dagger$; therefore, $J' = J^*$, $\mathbf{R}' = \mathbf{R}^{-1}$, and $\mathbf{B}' = \mathbf{B}$.

[6] If the Jones matrix is defined as a pair of columnn vectors, then the Gram matrix of the columns of $\mathbf{J}'$ would be used to determine its Hermitian component.

The unknown factor of $T = IG^2$ can be eliminated by noting that $|\mathbf{B}| = 1$; therefore, $|\boldsymbol{\rho}'| = T^2$ and $\mathbf{B} = (\boldsymbol{\rho}'/T)^{1/2}$.

4.7. *Axis-Angle Representation*

As shown in Appendix C.4, $\mathbf{B}$ and $\mathbf{R}$ can be expressed using axis-angle representation,

$$\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta) \equiv e^{\beta \underline{\hat{\boldsymbol{m}}}\cdot\underline{\boldsymbol{\sigma}}} = \boldsymbol{\sigma}_0 \cosh\beta + \underline{\hat{\boldsymbol{m}}}\cdot\underline{\boldsymbol{\sigma}} \sinh\beta \tag{59}$$

$$\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi) \equiv e^{i\phi \underline{\hat{\boldsymbol{n}}}\cdot\underline{\boldsymbol{\sigma}}} = \boldsymbol{\sigma}_0 \cos\phi + i\underline{\hat{\boldsymbol{n}}}\cdot\underline{\boldsymbol{\sigma}} \sin\phi, \tag{60}$$

where $\hat{\boldsymbol{m}}$ and $\hat{\boldsymbol{n}}$ are three-dimensional unit vectors, β and ϕ are real-valued scalars, and $\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3)$ is a 3-vector whose components are the Pauli spin matrices. Each axis-angle representation has 3 dof in the vector that defines it. Combined with the 2 dof in the real and imaginary parts of the complex-valued scalar J, Equation (56) has 8 dof, as expected for a 2×2 matrix with independent real and imaginary parts.

The axes and angles in Equations (59) and (60) have geometric interpretations in the four-dimensional space of the Stokes parameters. After a congruence transformation by a Hermitian matrix $\mathbf{B}$, the associated Stokes four-vector is transformed by a Lorentz boost along the $\hat{\boldsymbol{m}}$ axis by a hyperbolic angle -2β (e.g., Appendix D.1). A Lorentz transformation of the Stokes four-vector mixes S_0 with the polarization vector $\underline{\boldsymbol{S}}$, thereby altering both the intensity and the degree of polarization. Consequently, energy is not conserved. Hermititan matrices are equivalent to the *polconversion* defined by Hamaker (2000); they describe the diattenuation of a system (Lu & Chipman 1996) due to differential gain and non-orthogonality of the receptors.

In contrast, congruence transformation by a unitary matrix rotates the Stokes polarization vector $\underline{\boldsymbol{S}}$ about the $\hat{\boldsymbol{n}}$ axis by an angle -2ϕ, using the right-hand rule for rotation (e.g., Appendix D.2). This rotation in three-dimensional space leaves the total intensity and the degree of polarization unchanged. Unitary matrices are equivalent to the *polrotation* defined by Hamaker (2000); they describe the retardance of a system (Lu & Chipman 1996) due to differential phase and they represent any change of basis by projection onto a pair of orthonormal receptors (see Proof C.1).

The axis-angle representations of $\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)$ and $\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)$ exhibit a number of properties that prove useful during analysis. Both transformations are unimodular.

Proof 4.9.

$$\begin{aligned} |\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)| &= |\boldsymbol{\sigma}_0 \cosh\beta + \underline{\hat{\boldsymbol{m}}}\cdot\underline{\boldsymbol{\sigma}} \sinh\beta| && \text{Eqn. (59)} \\ &= \cosh^2\beta - |\underline{\hat{\boldsymbol{m}}}|^2 \sinh^2\beta && \text{Eqn. (C6)} \\ &= 1 \end{aligned}$$

Similarly, $|\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)| = \cos^2\phi - |i\underline{\hat{\boldsymbol{n}}}|^2 \sin^2\phi = 1$

Therefore, congruence transformation by $\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)$ or $\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)$ preserves the determinant of the coherency matrix and the invariant interval of the Stokes parameters.

Proof 4.10. Let $|\mathbf{U}| = 1$ and

$$\begin{aligned} |\boldsymbol{\rho}'| &= |\mathbf{U}\boldsymbol{\rho}\mathbf{U}^\dagger| && \text{Eqn. (37)} \\ &= |\mathbf{U}||\boldsymbol{\rho}||\mathbf{U}^\dagger| && |\mathbf{AB}| = |\mathbf{A}||\mathbf{B}| \\ &= |\boldsymbol{\rho}| && |\mathbf{U}| = |\mathbf{U}|^\dagger = 1 \end{aligned}$$

Exponentiation of $\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)$ or $\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)$ is equivalent to multiplying the angle by the power,

$$\mathbf{B}^y_{\underline{\hat{\boldsymbol{m}}}}(\beta) = \left(e^{\beta \underline{\hat{\boldsymbol{m}}}\cdot\underline{\boldsymbol{\sigma}}}\right)^y = e^{y\beta \underline{\hat{\boldsymbol{m}}}\cdot\underline{\boldsymbol{\sigma}}} = \mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(y\beta) \tag{61}$$

$$\mathbf{R}^y_{\underline{\hat{\boldsymbol{n}}}}(\phi) = \left(e^{i\phi \underline{\hat{\boldsymbol{n}}}\cdot\underline{\boldsymbol{\sigma}}}\right)^y = e^{iy\phi \underline{\hat{\boldsymbol{n}}}\cdot\underline{\boldsymbol{\sigma}}} = \mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(y\phi) \tag{62}$$

which simplifies the derivation of quantities such as the inverse or the Hermitian square root. Further conceptual benefits of the axis-angle representation are discussed in more detail in Appendix C.5.

5. SIGNAL PATH

In this section, each of the elements in the signal path depicted in Figure 4 are discussed in more detail, beginning with the phase convention $\boldsymbol{\Phi}$ and ending with Faraday rotation $\mathbf{F}$. The receiver gains, $\mathbf{G}$, and the deviations from an ideal feed, $\mathbf{D}$, are analyzed using the polar decomposition and expressed using the axis-angle representation. Apart from $\boldsymbol{\Phi}$, the remaining matrices are purely unitary (rotation) matrices.

5.1. *Phase Convention*

The phase convention of an instrument is defined by the number of signal processing steps that result in complex conjugation of the analytic signal, which cannot be represented by a linear transformation of the electric field vector, such as a Jones matrix.

Proof 5.1. If $f(z) = z^*$ and c and z are two complex numbers,

$$f(cz) = (cz)^* = c^* z^* \neq c f(z);$$

therefore, complex conjugation is not a linear mapping, which must satisfy $f(cz) = c f(z)$.

Referring to Equations (18) through (20), complex conjugation of the electric field vector negates the sign of S_3, a transformation represented by the impure Mueller matrix $\underline{\underline{\mathbf{\Phi}}}$ that appears in Equation (42). Consequently, the measured sign of S_3 is impacted by the treatment of the complex phase of the electric field, which depends on the phase convention adopted during the design and implementation of instrumentation and signal processing software; the method used to down-convert the radio signal; and the Nyquist zone that is digitally sampled.

As noted in Section 2 of van Straten et al. (2010), many signal processing textbooks and software packages adopt the convention that the argument of a complex-valued wave increases linearly with time, such that $z(t) \propto \exp(i\omega t)$. This convention is adopted in equation (1) of Stokes (1852) and in the analytic representation of a monochromatic wave presented in Equation (5). It is opposite to the convention used in many physics text books (e.g., Born & Wolf 1980), where the solution for a plane-propagating monochromatic electromagnetic wave is presented as $\boldsymbol{e}(x,t) = \boldsymbol{e}_0 \exp(i[kx - \omega t])$.

Choice of phase convention also arises in the implementation of the Discrete Fourier Transform (DFT). For example, relative to Equation (A2), which is consistent with the definition adopted by Bracewell (1965) and the implementation of commonly used DFT libraries such as FFTW (Frigo & Johnson 2005), equations (12.1.7) and (12.1.9) of Numerical Recipes (Press et al. 1986) define the DFT and its inverse using the opposite sign for the argument of the complex exponential. Application of a DFT that is based on an opposite sign convention is equivalent to negating the frequency, which for a real-valued input signal results in complex conjugation.

The configuration of observatory instrumentation can also negate phase. As shown in Appendix G, both complex conjugation and negation of frequency occur when either lower-sideband down-conversion is used; or, during dual-sideband down-conversion, the quadrature component is mixed with a local oscillator that *leads* that of the in-phase component by 90°; or an even-numbered Nyquist zone is sampled. Complex conjugation at any stage of analysis negates phase in all subsequent processing stages.

5.2. *Receiver Gains*

A radio receiver converts an electromagnetic wave travelling in free space into two separate signals, each representing an orthogonal component of the electric field vector. These separate signals propagate along distinct transmission lines and are independently processed using components such as amplifiers/attenuators, splitters/couplers, and analog-to-digital converters. Assuming that there is no further cross-coupling between the two signals, the receiver gains matrix

$$\mathbf{G} = \begin{pmatrix} G_0 & 0 \\ 0 & G_1 \end{pmatrix}, \tag{63}$$

describes the complex-valued gains, G_k, of all components used to process the two separate signals.

Intuition might suggest that it would be fruitful to represent the complex-valued gains in polar form,

$$G_k = g_k \exp i\phi_k,$$

and decompose $\mathbf{G}$ into separate amplitude and phase terms,

$$\mathbf{G} = \begin{pmatrix} g_0 & 0 \\ 0 & g_1 \end{pmatrix} \begin{pmatrix} e^{i\phi_0} & 0 \\ 0 & e^{i\phi_1} \end{pmatrix}. \tag{64}$$

However, when a promising next step is not obvious, it pays to employ a systematic approach. The strategy adopted in this work is based on the polar decomposition (Eqn. 56), the axis-angle representations of the boost and rotation components (Eqns. [59] and [60]), and determination of the boost component via the Gram matrix (Eqn. 57).

The polar decomposition, $\mathbf{G} = G\,\mathbf{B}_g\mathbf{R}_g$, includes the absolute gain,

$$G = |\mathbf{G}|^{1/2} = (G_0 G_1)^{1/2}, \tag{65}$$

a boost component $\mathbf{B}_g$, and a rotation component $\mathbf{R}_g$. As noted in Section 4.6, the absolute phase is chosen such that G is real-valued. The following sub-sections focus on the axis-angle representations of $\mathbf{B}_g$ and $\mathbf{R}_g$, where it is shown that $\mathbf{B}_g$ characterizes the ratio of the gain amplitudes, or differential gain, and $\mathbf{R}_g$ describes the differential phase.

5.2.1. *Differential Gain*

Using Equation (57), the boost component of $\mathbf{G}$ can be derived via the Gram matrix of its rows,

$$\mathbf{B}_g^2 = G^{-2}\,\mathbf{G}\mathbf{G}^\dagger = G^{-2} \begin{pmatrix} g_0^2 & 0 \\ 0 & g_1^2 \end{pmatrix}, \tag{66}$$

where $g_k = |G_k|$ are the amplitudes of the complex gains, such that $G = (g_0 g_1)^{1/2}$. Taking the square root

of both sides of Equation (66) yields

$$\mathbf{B}_g = \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma^{-1} \end{pmatrix}, \tag{67}$$

where $\Gamma = (g_0/g_1)^{\frac{1}{2}}$ is the real-valued amplitude ratio that describes the differential gain.

To arrive at the axis-angle representation of $\mathbf{B}_g$, note that its off-diagonal elements are zero and only $\boldsymbol{\sigma}_0$ and $\boldsymbol{\sigma}_1$ contribute to the diagonal elements of the sum in Equation (59). Therefore, $\underline{\hat{\boldsymbol{m}}} = (1,0,0)^T$ and

$$\mathbf{B}_g = \boldsymbol{\sigma}_0 \cosh\gamma + \boldsymbol{\sigma}_1 \sinh\gamma = \begin{pmatrix} e^{\gamma} & 0 \\ 0 & e^{-\gamma} \end{pmatrix}. \tag{68}$$

Equating the right hand sides of Equations (67) and (68) yields

$$\gamma = \ln\Gamma = \frac{1}{2}\ln\frac{g_0}{g_1}. \tag{69}$$

Differential gain describes any transformation that subjects opposite polarizations to different levels of amplification, also known as *diattenuation*. It can vary as a function of both time and frequency for a variety of reasons. For example, to keep the signal power within the optimal regime of operation, some experiments will set new attenuation levels at the start of each observation. Some instruments employ active attenuators that introduce differential gain fluctuations on short timescales. Furthermore, the mismatched responses of the components in the signal path typically lead to variation of β as a function of frequency.

5.2.2. *Differential Phase*

The axis-angle representation of $\mathbf{R}_g$ is obtained by rearranging the polar decomposition of $\mathbf{G}$ to arrive at $\mathbf{R}_g = G^{-1}\mathbf{B}_g^{-1}\mathbf{G}$. Both $\mathbf{G}$ and $\mathbf{B}_g$ (and their inverses) are diagonal and, as a product of diagonal matrices, $\mathbf{R}_g$ is also diagonal. Only $\boldsymbol{\sigma}_0$ and $\boldsymbol{\sigma}_1$ contribute to the diagonal elements of the sum in Equation (60); therefore, $\underline{\hat{\boldsymbol{n}}} = (1,0,0)^T$ and

$$\mathbf{R}_g = \boldsymbol{\sigma}_0 \cos\phi + i\boldsymbol{\sigma}_1 \sin\phi = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}, \tag{70}$$

where ϕ describes the differential phase. Subtracting the (degenerate) absolute phase $\phi_{\rm abs} = (\phi_0+\phi_1)/2$ from the phases in Equation (64) leaves $\phi = (\phi_0 - \phi_1)/2$.

Differential phase describes any transformation that subjects opposite polarizations to different delays; it is also known as *retardance* in optics and *cross-hand phase* or delay in interferometry. It arises when the opposite polarizations propagate along signal paths of different lengths and/or with different speeds. For a given delay Δt between $e_0(t)$ and $e_1(t)$, the corresponding differential phase $\phi = \nu\Delta t$ varies linearly with radio frequency. Nonlinear spectral variations arise from dispersive effects in the electronics of the receiver, down-conversion system, cables, and other components used to connect observatory instrumentation. Section 5.6 discusses the differential phase that arises during propagation through the magnetized plasma in the Earth's ionosphere and the interstellar medium.

5.3. *Deviations from an Ideal Feed*

In an ideal feed, the receptors have maximal response to orthogonal senses of either linear or circular polarization. However, in practice, the receptors are neither perfectly orthogonal nor exactly linearly- or circularly-polarized. Let non-ideal receptors be represented by unit row vectors, $\hat{\boldsymbol{r}}_0$ and $\hat{\boldsymbol{r}}_1$ ($\hat{\boldsymbol{r}}_i\hat{\boldsymbol{r}}_i^{\dagger} = 1$; note that the gains of the receptors are modeled by G_0 and G_1 in the previous section). Therefore, the deviations from an ideal feed are given by

$$\mathbf{D} = \begin{pmatrix} \hat{\boldsymbol{r}}_0 \\ \hat{\boldsymbol{r}}_1 \end{pmatrix} = \begin{pmatrix} \hat{r}_{00} & \hat{r}_{01} \\ \hat{r}_{10} & \hat{r}_{11} \end{pmatrix}. \tag{71}$$

Deviations from ideal are also known as *cross-talk*, *leakage*, or simply *D-terms*. The polar decomposition, $\mathbf{D} = G_d\,\mathbf{B}_d\mathbf{R}_d$, includes a scalar gain G_d, boost $\mathbf{B}_d$, and rotation $\mathbf{R}_d$. Rotation matrices describe only orthonormal pairs of receptors (see Proof C.1); therefore, the orthonormal component of $\mathbf{D}$ is given by $\mathbf{R}_d$ and any non-orthogonality must be described by $\mathbf{B}_d$, as detailed in the following sections.

5.3.1. *Non-orthogonal Component*

Using Equation (57), the boost component of $\mathbf{D}$ can be derived via the Gram matrix of its rows,

$$\mathbf{B}_d^2 = G_d^{-2}\mathbf{D}\mathbf{D}^{\dagger} = G_d^{-2}\begin{pmatrix} 1 & \hat{\boldsymbol{r}}_0\cdot\hat{\boldsymbol{r}}_1^{\dagger} \\ \hat{\boldsymbol{r}}_1\cdot\hat{\boldsymbol{r}}_0^{\dagger} & 1 \end{pmatrix}, \tag{72}$$

Let $\mathbf{B}_d^2 = \mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(2\beta)$ and note that the elements on the diagonal of Equation (72) are equal to each other; therefore, there can be no contribution from $\boldsymbol{\sigma}_1$. Only $\boldsymbol{\sigma}_2$ and $\boldsymbol{\sigma}_3$ contribute to the off-diagonal elements of the sum in Equation (59); i.e., $\underline{\hat{\boldsymbol{m}}} = (0, m_2, m_3)^T$, and

$$\begin{aligned}\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(2\beta) &= \boldsymbol{\sigma}_0\cosh 2\beta + (m_2\boldsymbol{\sigma}_2 + m_3\boldsymbol{\sigma}_3)\sinh 2\beta \\ &= \begin{pmatrix} \cosh 2\beta & z_m^*\sinh 2\beta \\ z_m\sinh 2\beta & \cosh 2\beta \end{pmatrix},\end{aligned} \tag{73}$$

where $z_m = m_2 + im_3$. Equating the diagonal elements of Equations (72) and (73) yields $G_d^2 = \operatorname{sech} 2\beta$; there-

fore, after taking the determinant of both sides of Equation (57) (noting that $\mathbf{B}_{\underline{\hat{m}}}(2\beta)$ is unimodular) and rearranging,

$$|\mathbf{D}\mathbf{D}^\dagger| = G_d^4 = \operatorname{sech}^2 2\beta = 1 - \tanh^2 2\beta. \quad (74)$$

Similarly, the determinant of Equation (72) leads to

$$|\mathbf{D}\mathbf{D}^\dagger| = 1 - |\hat{\boldsymbol{r}}_0 \cdot \hat{\boldsymbol{r}}_1^\dagger|^2. \quad (75)$$

Equating the right-hand sides of Equations (74) and (75) yields

$$\beta = \frac{1}{2} \tanh^{-1} \left| \hat{\boldsymbol{r}}_0 \cdot \hat{\boldsymbol{r}}_1^\dagger \right|. \quad (76)$$

When the receptors are orthogonal, β is zero and $\mathbf{B}_d$ reduces to the identity matrix (no deviations from ideal).

As in Section 4.1 of van Straten (2004), the 2 dof in the non-orthogonal component of $\mathbf{D}$ may be parameterized by $\underline{\boldsymbol{b}} = (0, b_2, b_3)^T = \underline{\hat{\boldsymbol{m}}} \sinh\beta$, such that $\sinh^2\beta = \underline{\boldsymbol{b}} \cdot \underline{\boldsymbol{b}}$, $\cosh^2\beta = 1 + \underline{\boldsymbol{b}} \cdot \underline{\boldsymbol{b}}$, and

$$\mathbf{B}_d = \mathbf{B}_{\underline{\hat{m}}}(\beta) = \boldsymbol{\sigma}_0 (1 + \underline{\boldsymbol{b}} \cdot \underline{\boldsymbol{b}})^{1/2} + \underline{\boldsymbol{\sigma}} \cdot \underline{\boldsymbol{b}}. \quad (77)$$

To first order, this derivation is consistent with Equation (19) of Britton (2000), which describes non-orthogonal receptors using a product of two separate boosts along the S_2 and S_3 axes (corresponding to Stokes U and Stokes V, respectively, in a linear basis). In general, the result of this product includes a rotation about the S_1 (Stokes Q) axis; however, if either of these two boosts is small, then this rotation is negligible. In contrast, no small-value approximations are made in the derivation of Equation (77).

The instrumental boost due to non-orthogonal receptors can vary as a function of time and frequency. For example, the parallactic rotation of the receiver during the transit of the source changes the orientation of $\underline{\hat{\boldsymbol{m}}}$ with respect to the celestial coordinate system. Furthermore, electronic cross-coupling between the two receptors can vary with radio frequency.

5.3.2. *Orthonormal Component*

In principle, $\mathbf{R}_d$ can be obtained by rearranging the polar decomposition of $\mathbf{D}$ to arrive at $\mathbf{R}_d = G_d^{-1}\mathbf{B}_d^{-1}\mathbf{D}$; however, this proves to be algebraically cumbersome. Instead, following Britton (2000), the transformation is expressed using the spherical coordinates ψ and χ that define the polarization ellipse.

First, in the ideal basis, consider a 100% polarized signal $\boldsymbol{e}(t)$ with Stokes polarization vector (see Eqn. 7)

$$\underline{\boldsymbol{S}} = I \left(\cos 2\chi \cos 2\psi,\ \cos 2\chi \sin 2\psi,\ \sin 2\chi\right)^T. \quad (78)$$

Next, consider the transformation from the ideal basis to one in which the first receptor responds maximally to $\boldsymbol{e}(t)$ and the response in the orthogonal receptor is zero. In this basis, the measured coherency matrix,

$$\boldsymbol{\rho}' = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} = (\boldsymbol{\sigma}_0 + \boldsymbol{\sigma}_1)\, I/2 \quad (79)$$

and the polarization vector $\underline{\boldsymbol{S}}' = I(1,0,0)^T$.

With reference to Figure 2, $\underline{\boldsymbol{S}}'$ is obtained by rotating $\underline{\boldsymbol{S}}$ about the S_3 axis by -2ψ in the ideal basis, then rotating the result by 2χ about the S_2' axis in the intermediate basis. The corresponding transformation of the electric field, $\boldsymbol{e}'(t) = \mathbf{R}_d\, \boldsymbol{e}(t)$, where

$$\mathbf{R}_d = \mathbf{R}_{\underline{\hat{\boldsymbol{2}}}}(-\chi)\mathbf{R}_{\underline{\hat{\boldsymbol{3}}}}(\psi). \quad (80)$$

This is equivalent to equation (15) of Britton (2000), apart from the negation of χ that associates positive values of χ with $S_3 > 0$.

This is a useful parameterization of the 2 dof in $\mathbf{R}_d$ because it reduces to the identity matrix (no deviations from ideal) when χ and ψ are zero. Furthermore, because differential phase results in a rotation about the S_1 axis, it is desirable to describe the deviation from the ideal basis transformation using rotations about the other two axes. Together, the three rotations define the orientation of the basis in a manner similar to the Tait–Bryan angles (yaw, pitch, and roll).

5.4. *Nominal Feed Configuration*

As in van Straten et al. (2010), the nominal feed configuration is described by three parameters that are summarized in Table 1. The *feed basis* defines the polarizations of the receptors (linear or circular); the *feed hand* defines the handedness of the basis (left or right); and the *symmetry angle* describes the orientation of the reference frame defined by the receptors. The corresponding basis transformation

$$\mathbf{C} = \mathbf{R}_h \mathbf{R}_b \mathbf{R}_{\underline{\hat{\boldsymbol{z}}}}(\Theta) \quad (81)$$

is the product of feed hand $\mathbf{R}_h$ and basis $\mathbf{R}_b$ transformations, and a rotation about the line of sight $\mathbf{R}_{\underline{\hat{\boldsymbol{z}}}}(\Theta)$ defined by the symmetry angle Θ. These are detailed in the following sections, starting with the feed basis, which defines both the nominal symmetry angle and the effect of the feed hand transformation.

5.4.1. *Feed Basis*

Typically, the two receptors in a receiver are either linearly or circularly polarized, and Section 2 describes

Table 1. The effects of phase convention and nominal feed configuration parameters in each basis.

Parameter	Range	Effect	
		Linear	Circular
Phase Convention	$\pm$ 1	$\pm$ V	$\pm$ U
Feed Basis	LIN or CIRC	$\underline{\boldsymbol{S}} = (Q, U, V)^T$	$\underline{\boldsymbol{S}} = (V, Q, U)^T$
Feed Hand	$\pm$ 1	$\pm$ Q&V	$\pm$ U&V
Symmetry Angle	$-\pi/2 < \theta < \pi/2$	$\mathbf{R}_{\hat{\underline{\boldsymbol{z}}}}(\theta - \pi/4)$	$\mathbf{R}_{\hat{\underline{\boldsymbol{z}}}}(\theta)$

a Cartesian coordinate system defined by linearly polarized receptors. As shown in Appendix D.4, the transformation from this Cartesian basis to one defined by a pair of circularly-polarized receptors is described by

$$\mathbf{R}_b = \frac{1}{\sqrt{2i}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}. \tag{82}$$

In the basis defined by $\mathbf{R}_b$, the Stokes polarization vector $\underline{\boldsymbol{S}} = (S_1, S_2, S_3)^T = (V, Q, U)^T$ and the effects of all subsequent transformations (from $\mathbf{X}$ to $\underline{\underline{\boldsymbol{\Phi}}}$) differ from their impact in the original Cartesian basis.

For example, in the circular basis, the rotations around the S_2 and S_3 axes that define both the polarization ellipse (Eqn. 4) and the orthonormal component of deviation (Eqn. 80) are better described by pair of polar angles, ξ and ζ, respectively. These angles, depicted in Figure 5, are related to the orientation and ellipticity angles by

$$\tan 2\xi = \cot 2\chi \cos 2\psi \tag{83}$$

$$\sin 2\zeta = \cos 2\chi \sin 2\psi. \tag{84}$$

In the circular basis, the differential phase ϕ describes rotation about the Stokes V axis, which is equivalent to a physical rotation about the line of sight (Proof 5.2). Later sections describe the impacts of the circular basis on the feed hand $\mathbf{R}_h$ (Section 5.4.2), axis of symmetry $\mathbf{R}_{\hat{\underline{\boldsymbol{z}}}}(\Theta)$ (Section 5.4.3), and phase convention $\underline{\underline{\boldsymbol{\Phi}}}$ (Section 6.1) transformations.

5.4.2. *Feed Hand / Basis Reflection*

The feed hand is determined by the design and integration of the receiver, such as the sign convention for the voltages output by an orthomode transducer, which determines the directions of the x and y axes in Figure 1; or any cable swaps that might occur between the receiver and the backend, which swap the x and y axes. Feed hand is also impacted by the total number of reflections in the antenna structure, which depends upon the placement of the feed at the primary or secondary focus, and the number of reflections in any beam waveguide structure (e.g., Grueff et al. 2004; Petrov et al. 2015) or corrective mirrors designed to optimize sensitivity (e.g., Kildal et al. 1994; Granet & James 1997).

Normal reflection by a conducting surface, such that the angles of incidence and reflection equal zero and the incident and reflected rays are anti-parallel, can be conceptualized in two different ways that arrive at the same result. One way to model the observation of reflected rays is by turning over the feed horn, which originally points up at the sky, to point down at the reflector. This approach is detailed in Appendix H.

Alternatively, the change in wave direction is modeled by negating the z-axis, thereby producing a left-handed coordinate system. The handedness of the reference frame is also negated by exchanging the components of the electric field. With reference to Equations (18) through (20), swapping e_x and e_y negates Stokes Q and V, which is equivalent to a $\pm 180°$ rotation about the Stokes U axis. In a basis defined by circularly-polarized receptors, swapping e_L and e_R reverses the signs of both Stokes U and V, which is equivalent to a $\pm 180°$ rotation about the Stokes Q axis. In both cases, the rotation is given by $\mathbf{R} = \boldsymbol{\sigma}_2$; it negates both the ellipticity χ (Eqn. 10) and the position angle ψ (Eqn. 9).

Although the number of reflections in an antenna can be easily counted, it is generally less feasible to account for voltage negation or cable swapping. Therefore, the feed hand is typically determined experimentally with reference to previously published polarization data. For example, as both χ and ψ are negated upon reflection, the feed hand can be unambiguously determined by observing a source with known ellipticity or Faraday rotation measure. Given the total number of hand reversals N, the feed hand matrix is defined by

$$\mathbf{R}_h = \boldsymbol{\sigma}_2^N. \tag{85}$$

The negation of the position angle after normal reflection from a metal surface is noted in the context of a more richly detailed treatment of the physics of reflection and refraction by Born & Wolf (1980). However, it appears to be unrecognized in some important works on radio polarimetry, which mention only the negation of Stokes V (e.g., Hamaker et al. 1996; Robishaw & Heiles 2021). Appendix A of Clarke (2010) provides a carefully detailed derivation of both the linear and circular

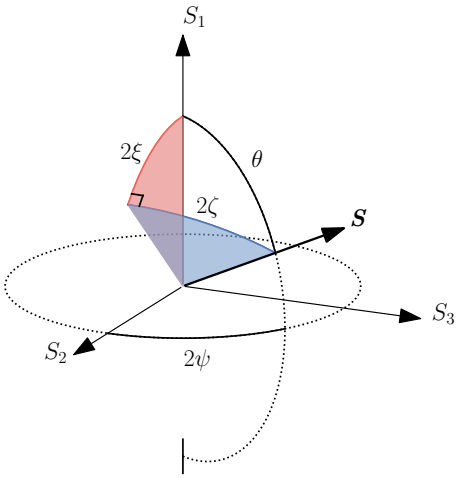

Figure 5. In the circular basis, the direction of the polarization vector, $\underline{\boldsymbol{S}} = (S_1, S_2, S_3)^T = (V, Q, U)^T$, is defined by the polar angles, ξ and ζ. These angles together define a right triangle on the surface of the Poincaré sphere with hypotenuse equal to the colatitude, $\theta = \pi/2 - 2\chi$, measured with respect to the pole defined by positive S_1 (Stokes V). Also depicted is the longitude 2ψ, measured with respect to positive S_2 (Stokes Q) toward positive S_3 (Stokes U).

polarization of radiation that is observed after reflection, including a historical account of some confusion that has persisted on this topic.

5.4.3. *Axis of Symmetry*

The axis of symmetry in the x-y plane depicted in Figure 1 is defined as the position angle of a linearly-polarized wave that produces equal responses in each ideal receptor. In a basis defined by linearly-polarized receptors, a linearly-polarized wave that oscillates along $x = y$ (positive Stokes U) will produce equal and in-phase responses in each receptor; therefore, the symmetry angle has a nominal value of 45°. In a basis defined by circularly-polarized receptors, a linearly-polarized wave that oscillates along the x axis will produce equal and in-phase responses in each receptor, and the symmetry angle has a nominal value of 0°.

The symmetry axis of a receiver on steerable mount describes its rotation about the line of sight with respect to a reference point on the antenna structure. For a fixed dipole on the ground, the symmetry axis describes its rotation about the zenith. Once established, the symmetry angle is typically treated as a constant. Any unintended rotation (either geometric or apparent; e.g., owing to cross-coupling of the receptors) must be determined through calibration of the orthonormal component of deviation, which includes an orientation angle ψ (Eqn. 80). Other known rotations are included in the projection transformation described in Section 5.5.

5.5. *Projection onto the Celestial Sphere*

For radio astronomical observations, the incident electric field measured in the reference frame of the antenna must be transformed to the celestial reference frame adopted by the IAU. In this coordinate system, the electric field vector is described by its projection onto the plane of the sky, with basis vectors $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ pointing North and East, respectively. When observing with a fully-steerable antenna, this transformation is described by a rotation about the line of sight through the parallactic angle, which is well determined by geometry. In the PSRCHIVE software (Hotan et al. 2004), the transformation to celestial coordinates can be computed for a wide variety of antenna mounts, or it can optionally be replaced by a user-supplied table of transformations that describe the direction-dependent response of the antenna, as in the case of a dipole array (e.g., Asad et al. 2016; Sokolowski et al. 2017).

Without a detailed electromagnetic model of a fixed dipole, the direction-dependent response can be described to first order as a geometric projection of the receptors onto the sky. Such a transformation is a product of a rotation about the Stokes V axis that describes parallactic rotation about the line of sight, and a boost along an axis in the Stokes $Q - U$ plane that describes the combined effects of differential gain and apparent nonorthogonality due to foreshortening of the projected receptors. The combination of these effects causes the symmetry axis (see Appendix C.5) of the projection transformation to vary with the direction to the source.

Therefore, for a fixed dipole array, there is no fundamental degeneracy as described in Appendix I.1, which is a consequence of a single constant axis of symmetry. That is, in principle, it is possible to determine 6 dof of the response (all but the absolute gain) by observing a polarized point source with a phased array of fixed dipoles over a wide range of hour angles. However, as described in Appendix I.2, near multicollinearity may arise when the rapidity of the boost along the Stokes V axis is allowed to vary, causing numerical instability and inflated uncertainty of the rapidity estimate.

The Lorentz boost caused by projection of fixed dipoles onto the celestial sphere converts significant levels of unpolarized flux into linearly-polarized flux. In wide-field interferometric imaging experiments, this can be exploited to fully constrain the instrumental response using the unpolarized sky (e.g., Byrne et al. 2022; Kansabanik et al. 2025). A similar approach is generally not possible in single-antenna observations owing to a combination of stray radiation (e.g. Lockman 2002), radio frequency interference (e.g. Offringa et al. 2010), and the

unknown variation of the instrumental response over the beam (Eqn. 55).

5.6. *Faraday Rotation*

During propagation through the interstellar medium or the Earth's ionosphere, an astrophysical signal experiences Faraday rotation. This phenomenon, also known as circular birefringence, arises because the natural modes of propagation in a magnetized (non-relativsitic and collisionless) plasma are circularly polarized, such that LCP and RCP components of the electric field propagate with different speeds. At a given frequency, the resulting differential delay between LCP and RCP components causes the electric field to rotate about the line of sight. For example, if the signal propagates in the direction of the magnetic field vector, then LCP is delayed with respect to RCP (Manchester 1972; Ferrière et al. 2021); the resulting transformation is equivalent to rotating the electric field by $\Delta\Psi$ around the line of sight (counter-clockwise as seen by the observer).

Proof 5.2. In the circular basis (Eqn. 82), apply differential phase (Eqn. 70), and return to the original Cartesian basis.

$$\begin{aligned}\mathbf{F}(\Delta\Psi) &= \mathbf{R}_b^{-1}\mathbf{R}_g\mathbf{R}_b \\ &= \frac{1}{2}\begin{pmatrix}1 & 1\\ i & -i\end{pmatrix}\begin{pmatrix}e^{-i\Delta\Psi} & 0\\ 0 & e^{i\Delta\Psi}\end{pmatrix}\begin{pmatrix}1 & -i\\ 1 & i\end{pmatrix} \\ &= \begin{pmatrix}\cos\Delta\Psi & -\sin\Delta\Psi\\ \sin\Delta\Psi & \cos\Delta\Psi\end{pmatrix} \\ &= \mathbf{R}_{\underline{\hat{\boldsymbol{z}}}}(-\Delta\Psi)\end{aligned}$$

The change in the position angle of the radiation,

$$\Delta\Psi(\nu) = \mathrm{RM}\, c^2\nu^{-2}, \tag{86}$$

where RM is the rotation measure, c is the speed of light, and ν is the frequency of the radiation. Note that birefringence is dispersive and the phase shift / rotation is proportional to ν^{-2}; this differs from the differential phase due to a path length difference, which is proportional to ν. The RM is proportional to the path integral of the density of free electrons, n_e, times the strength of the magnetic field, $\underline{\boldsymbol{B}}$, parallel to the direction of propagation $d\underline{\boldsymbol{z}}$ (e.g., Han et al. 2006)

$$\mathrm{RM} = C\int_0^D n_e\, \underline{\boldsymbol{B}}\cdot d\underline{\boldsymbol{z}}, \tag{87}$$

where D is the distance to the astrophysical source and

$$C = 0.81\frac{\mathrm{rad\,m^{-2}}}{\mathrm{cm^{-3}}\,\mu\mathrm{G\,pc}}. \tag{88}$$

If the magnetic or ionic properties of the ionosphere fluctuate rapidly, either on timescales shorter than the interval over which the Stokes parameters are integrated or on spatial scales smaller than the volume sampled by multipath propagation (e.g., Cordes et al. 2016; Suresh & Cordes 2019), then the signal will be depolarized by stochastic Faraday rotation (Melrose & Macquart 1998), resulting in an impure Mueller matrix with no equivalent Jones matrix (see Appendix C.7).

6. CALIBRATION

The response of each element of the signal path may include an unknown component that must be determined experimentally. This section briefly reviews methods of calibrating each component in Equation (42) and Figure 4, focusing primarily on the response of a single antenna (single dish or phased array). Where relevant, additional notes on interferometric calibration are included.

6.1. *Phase Convention*

For linearly-polarized receptors, S_3 equals Stokes V, and complex conjugation negates the ellipticity angle. For circularly-polarized receptors, S_3 equals Stokes U, and the position angle is negated by conjugation. Complex conjugation of the electric field vector also negates frequency in the spectral domain. Therefore, the phase convention used to represent the electric field can be determined by incorporating information from other sources. For example, owing to dispersion, frequency negation is immediately apparent during radio pulsar signal analysis, and the phase convention must be known to perform phase-coherent dispersion removal in radio pulsar observations (e.g., van Straten & Bailes 2011). Frequency negation can also be detected in continuum observations over regions of the radio spectrum that include strong, well-defined spectral lines. In an interferometer, complex conjugation of the visibility matrix computed for each antenna pair results in a 180° rotation of the image of the brightness distribution on the plane of the sky.

6.2. *Receiver Gains*

The receiver gains matrix, $\mathbf{G}$, is often determined using observations of an artificial reference source with a well-defined polarization state. For example, many receivers have a built-in source of broadband noise (e.g., a noise diode coupled to the receptors) that can be used to estimate gains as a function of radio frequency. Determination of $\mathbf{G}$ using only an artificial noise source is based on the *ideal feed assumptions*, in which the receptors are

orthogonal and there is no cross-coupling between them (i.e., $\mathbf{D}$ is the identity matrix); and the artificial reference source illuminates the receptors equally (i.e., the receptors have identical responses to the reference source, in both amplitude and phase). Although these assumptions are common, it is more accurate to experimentally determine both $\mathbf{D}$ and the Stokes parameters intrinsic to the artificial reference source (e.g., van Straten 2004), as described in Section 6.3.

Even when fully modeling the deviations from an ideal feed and polarization of the artificial reference source, the ideal feed assumptions can provide a useful, and sometimes necessary, first guess for $\mathbf{G}$. Therefore, they are an important part of polarimetric calibration for both single antennas (e.g., von Hoensbroech & Xilouris 1997; Navarro et al. 1997; Gould & Lyne 1998; Han et al. 2009) and interferometric arrays (e.g., Johnston et al. 2008; Serylak et al. 2021). As shown in Section 5.2, the polar decomposition of the complex gains,

$$\mathbf{G} = G\,\mathbf{B}_g(\gamma)\mathbf{R}_g(\phi), \tag{89}$$

has 3 dof: the absolute gain G, the differential gain γ, and the differential phase ϕ. To experimentally determine G requires a *standard candle*, i.e., a source with known flux density at the radio frequency of the observations. Absolute gain calibration (also known as flux calibration) is described in more detail in Section 7.2 of van Straten et al. (2012). In an interferometer, unique values of the complex gains must be determined for each element in the array.

Differential gain mixes only S_0 and S_1, and differential phase mixes only S_2 and S_3. Therefore, given observations of an ideal artificial reference source with known polarization, the two unknown values of γ and ϕ can be solved independently. The following two subsections derive solutions for γ and ϕ based on the assumption that a pure linearly polarized reference source induces identical signals in the orthogonal receptors of an ideal feed. In the linear basis, the Stokes parameters intrinsic to such a source,

$$\underline{\underline{\boldsymbol{S}}}_{\rm ref} = I_{\rm ref}[1,0,1,0]^T, \tag{90}$$

where $I_{\rm ref}$ is the intensity of the reference source.

6.2.1. *Differential Phase*

As a rotation about the S_1 axis, differential phase mixes only S_2 and S_3, and to determine ϕ requires observations of a polarized reference source with known intrinsic values of S_2 and S_3. In the linear basis, S_2 and S_3 correspond to Stokes U and Stokes V, respectively; in the circular basis, S_2 and S_3 correspond to Stokes Q and Stokes U. Given the measured Stokes parameters of the reference source, with observed values of S'_2 and S'_3, Equations (19) and (20) can be solved for the intrinsic and observed values of $\langle e_0^*(t)e_1(t)\rangle$; e.g.

$$\langle e_0^*(t)e_1(t)\rangle = (S_2 + iS_3)/2. \tag{91}$$

To relate the intrinsic values to their observed values, consider the transformation of the intrinsic electric field vector $\boldsymbol{e}(\mathrm{t})$ into the observed electric field vector as described by Equation (70),

$$\boldsymbol{e}'(t) = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix} \boldsymbol{e}(t). \tag{92}$$

The observed value of

$$\langle e_0'^*(t)e_1'(t)\rangle = e^{-2i\phi}\langle e_0^*(t)e_1(t)\rangle \tag{93}$$

can then be used to solve for ϕ; i.e.,

$$\phi = -\frac{1}{2}\arg\left(\frac{\langle e_0'^*(t)e_1'(t)\rangle}{\langle e_0^*(t)e_1(t)\rangle}\right), \tag{94}$$

where $\arg(z)$ is the argument of complex number z.

If the reference source produces an in-phase signal in each receptor, such that the differential phase between $e_0(t)$ and $e_1(t)$ equals zero, then $\langle e_0^*(t)e_1(t)\rangle$ is real-valued and

$$\phi = -\frac{1}{2}\arg\left(\langle e_0'^*(t)e_1'(t)\rangle\right) = -\frac{1}{2}\tan^{-1}(S'_3/S'_2). \tag{95}$$

6.2.2. *Differential Gain*

As a boost along the S_1 axis, differential gain mixes only S_0 and S_1; therefore, determination of γ requires observations of a polarized reference source with known intrinsic values of S_0 and S_1, where S_1 corresponds to Stokes Q in the linear basis and to Stokes V in the circular basis. Given the measured Stokes parameters of the reference source, with observed values of S'_0 and S'_1, Equations (17) and (18) can be solved for the intrinsic and observed values of $\langle|e_0(t)|^2\rangle$ and $\langle|e_1(t)|^2\rangle$; e.g.

$$\langle|e_0(t)|^2\rangle = (S_0 + S_1)/2 \tag{96}$$

$$\langle|e_1(t)|^2\rangle = (S_0 - S_1)/2. \tag{97}$$

To relate the intrinsic values to their observed values, consider the transformation of the intrinsic electric field vector $\boldsymbol{e}(\mathrm{t})$ into the observed electric field vector as described by Equation (67),

$$\boldsymbol{e}'(t) = G\begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma^{-1} \end{pmatrix} \boldsymbol{e}(t). \tag{98}$$

The observed values of

$$\begin{aligned}\langle|e'_0(t)|^2\rangle &= G^2\Gamma^2\langle|e_0(t)|^2\rangle \\ \langle|e'_1(t)|^2\rangle &= G^2\Gamma^{-2}\langle|e_1(t)|^2\rangle\end{aligned} \tag{99}$$

can be used to solve for Γ; i.e.,

$$\Gamma = \left(\frac{\langle|e'_0(t)|^2\rangle}{\langle|e'_1(t)|^2\rangle}\frac{\langle|e_1(t)|^2\rangle}{\langle|e_0(t)|^2\rangle}\right)^{1/4} \tag{100}$$

If the reference source produces a signal with equal power in each receptor, then $\langle|e_1(t)|^2\rangle = \langle|e_0(t)|^2\rangle$ and

$$\Gamma = \left(\frac{\langle|e'_0(t)|^2\rangle}{\langle|e'_1(t)|^2\rangle}\right)^{1/4} = \left(\frac{S'_0 + S'_1}{S'_0 - S'_1}\right)^{1/4}. \tag{101}$$

6.3. *Deviations from Ideal*

For a single antenna, the deviations from an ideal feed, $\mathbf{D}$, can be determined by modeling variations in the observed polarization of an astrophysical source as a function of parallactic angle (e.g., Stinebring et al. 1984; Xilouris 1991; McKinnon 1992; Heiles et al. 2001; Johnston 2002; van Straten 2004). In some treatments, $\mathbf{D}$ is estimated after the complex receiver gains, $\mathbf{G}$, have been calibrated; in others, $\mathbf{D}$ and $\mathbf{G}$ are jointly determined. Some of the models used to describe $\mathbf{D}$ are discussed and compared in Section 7. In an interferometer, unique values of $\mathbf{J}_i = \mathbf{G}_i\mathbf{D}_i$ must be determined for each element in the array, indexed by i. When used as a phased-array, only the Jones matrix of the sum, $\mathbf{J} = \sum_i \mathbf{J}_i$, must be determined.

More broadly, deviations from the ideal feed assumptions include the unintended polarization intrinsic to the artificial reference source, $\boldsymbol{\rho}_{\rm ref}$. If observations of the reference source are available, then $\boldsymbol{\rho}_{\rm ref}$ can be included in the model and jointly determined along with $\mathbf{G}$ and $\mathbf{D}$. When doing so, the reference source can be modeled as a free-space signal that is transmitted into the feed, or as a guided signal on a transmission line that is coupled after reception.

A free-space signal may be transmitted into the feed using a dipole with known orientation. Typically, the polarization of such an artificial reference signal is distorted by near-field effects, mutual coupling, resonances, and standing waves between the transmitting dipole, feed horn and antenna structure. When modeled as a free-space signal, the artificial reference source is transformed by the product, $\mathbf{GD}$, as in van Straten (2004).

Alternatively, the artificial reference source may be coupled to the astronomical signal after reception, at which point the two polarizations propagate as a pair of guided signals on separate transmission lines. This requires splitting the artificial reference source signal and coupling it identically to the two orthogonally-polarized signals, ideally before any amplification. Typically, the polarization of such a reference signal is distorted by mismatched impedance, insertion losses, reflections, and standing waves. When modeling an artificial reference source that is guided and coupled, $\boldsymbol{\rho}_{\rm ref}$ is transformed by only $\mathbf{G}$, as in Bailes et al. (2020).

In either case (free-space or guided transmission), the instrumental response to an artificial noise source typically differs from its response to an astrophysical source. However, in many experiments, the absolute polarization of the artificial reference source is irrelevant. For example, techniques used to update the differential gain and phase given known values of $\boldsymbol{\rho}_{\rm ref}$ and $\mathbf{D}$ (e.g., Ord et al. 2004) rely only on the temporal stability of these quantities. Therefore, any differences in the instrumental response to astrophysical and artificial sources can be absorbed by redefinition of $\boldsymbol{\rho}_{\rm ref}$.

This distinction is particularly important when using an astrophysical source as a polarized reference source (e.g., van Straten 2013). In this case, $\mathbf{D}$ may be corrupted by unmodeled ionospheric Faraday rotation (e.g., Rogers et al. 2024). To avoid also corrupting $\boldsymbol{\rho}_{\rm ref}$, it is necessary to model the artificial reference source as though it is coupled after $\mathbf{D}$, regardless of the actual receiver design.

6.4. *Nominal Feed Configuration*

Even the nominal feed configuration may require some reasoning, informed by observations, about the handedness and symmetry axis of the receiver. For example, any linear transformation of the electric field that negates either the ellipticity or the position angle must negate both. Therefore, regardless of the polarization of the receptors, if the sign of either the ellipticity or position angle of an observed source is known, then the handedness of the receptor basis can be unambiguously determined. Even without an estimate of the absolute value of the position angle, its sign can be inferred from either the RM or the slope of the canonical S-shaped sweep of position angle described by the Rotating Vector Model (RVM; Radhakrishnan & Cooke 1969) in longitude-resolved observations of the polarized emission from radio pulsars.

6.5. *Celestial Sphere Projection*

In principle, the projection between the celestial reference frame and the nominal receptor basis may be completely determined by known geometry; however, it may also require sophisticated modeling of direction-dependent effects. For example, gravitational deformation of an antenna (both the reflecting surface and the

support structure) can cause the polarimetric response to vary with pointing direction (Robishaw & Heiles 2021; Islam et al. 2024). Furthermore, the response of a dipole array varies with direction to the source in a manner that is not well-described by the first-order geometric effects of projection. Therefore, the accuracy with which the instrumental response can be determined will be limited by the fidelity of the electromagnetic model used to describe direction-dependent effects specific to the antenna design (e.g., Wijnholds et al. 2011; Mitchell et al. 2012; Sutinjo et al. 2015).

On some radio telescopes, either the entire reflector can be mechanically rotated about the line of sight (e.g., Hotan et al. 2021) or the receiver can be rotated about the line of sight with respect to the reflector (e.g., Staveley-Smith et al. 1996). This feature is typically used to compensate for the parallactic rotation of the observatory with respect to the sky; however, it can also be used to simulate observation over a range of parallactic angles (e.g., Guillemot et al. 2023). For such systems, the mechanical rotation can be included as a component of the projection matrix. For example, if the feed horn is rotated with respect to the reflector, the projection matrix may be decomposed as

$$\mathbf{P} = \mathbf{R}_{\underline{\hat{z}}}(\theta_{\rm feed})\mathbf{J}(l,m)\mathbf{R}_{\underline{\hat{z}}}(\theta_{\rm para}), \tag{102}$$

where $\theta_{\rm feed}$ is the feed horn rotation, $\mathbf{J}(l,m)$ models direction-dependent effects of the antenna (e.g., as a function of the direction cosines l and m that describe angular offsets from the primary axis), and $\theta_{\rm para}$ is the parallactic angle (e.g., Guillemot et al. 2025).

6.6. *Faraday Rotation*

To calibrate Faraday rotation, it is necessary to first estimate the Faraday Rotation Measure (RM). A wide variety of techniques have been developed for RM estimation, including directly modeling the variation in observed position angle $\Delta\Psi$ as a function of radio frequency (e.g., Noutsos et al. 2008); iteratively computing the weighted mean position-angle difference between two radio frequency bands (Han et al. 2006; Caleb et al. 2019); and searching for the peak in linearly-polarized flux after integrating over radio frequency as a function of trial RM (e.g., Hotan et al. 2005; Brentjens & de Bruyn 2005).

In principle, Faraday rotation can be corrected by rotating the Stokes polarization vector $\underline{\boldsymbol{S}}(\nu)$ observed at each frequency by $-\Delta\Psi(\nu)$ about the Stokes V axis. This effectively yields the position angle as though observed at infinite radio frequency. However, such a correction would later have to be inverted when refining the RM estimate, which requires a set of observations made at finite radio frequencies. Therefore, the Faraday rotation at each frequency ν is corrected with respect to the linear polarization state observed at some fiducial frequency ν_0 (typically the centre frequency of the band); i.e.,

$$\Delta\Psi'(\nu) = \mathrm{RM}\, c^2 \left(\nu^{-2} - \nu_0^{-2}\right). \tag{103}$$

The RM may also vary as a function of time over the duration of the integration. For example, the ionosphere can cause the RM to vary by several rad m^{-2} on timescales of hours. These variations can be predicted using a model of the geomagnetic field and a map of ionospheric electron content (e.g., Porayko et al. 2019). Temporal variations in RM also cause the position angle observed at the fiducial frequency to vary with time; therefore, Faraday rotation at time t is corrected by

$$\Delta\Psi(\nu,t) = \mathrm{RM}_0\, c^2 \left(\nu^{-2} - \nu_0^{-2}\right) + \Delta\mathrm{RM}(t)\, c^2\nu^{-2}. \tag{104}$$

where RM_0 is constant and $\Delta\mathrm{RM}(t) = \mathrm{RM}(t) - \mathrm{RM}_0$.

7. POLARIMETER MODEL SELECTION

Accurate polarimetric calibration requires a model of the instrumental response that includes deviations from an ideal feed. Several different approaches to decomposing the single-antenna instrumental response have been proposed in the published literature. Broadly, these models can be divided into those based on Jones matrices and those based on Mueller matrices; e.g., the Jones matrix equation (11) of Hamaker et al. (1996, hereafter HBS96); the Jones matrix equation (19) of Britton (2000, hereafter B2000); and the Mueller matrix equation (22) of Heiles et al. (2001, hereafter H2001). In principle, a Mueller matrix has an additional 9 dof and can describe any linear transformation of the Stokes parameters, including non-linear transformations of the electric field such as complex conjugation. However, these additional dof are not employed in H2001, where the Mueller matrices are derived from Jones matrices.

Regardless of the objects used to represent polarization state and transformations, when applied to experimental data analysis, a mathematical model of the instrumental response should be

1. physically motivated, such that model components describe elements of the signal path;
2. surjective, such that the parameter space spans all possible transformations of interest;
3. injective, such that each transformation is described by a unique set of parameters;

4. self-consistent, such that the assumed properties of the system are preserved; and

5. numerically stable, at least in the vicinity of the anticipated solution.

In the following subsections, these criteria are discussed in the context of single-antenna observations of a point source, primarily using the models proposed by HBS96, B2000, and H2001 as examples. When referring to the model parameters used in these works, the original mathematical symbols are retained; these may conflict with the symbols employed in the previous sections of this paper.

7.1. *Physical Motivation*

Hamaker (2000) proposed a purely algebraic decomposition of the instrumental response using a single polar decomposition, and this approach is one of the two implemented by van Straten (2004). Although simple and useful, the polar decomposition is not directly amenable to physically meaningful interpretation. It does not model the order in which elements in the signal chain are encountered, and it does not permit separation of backend and frontend transformations.

In practice, it proves useful to decompose the instrumental Jones matrix into a product of backend and frontend transformations, such as **G** and **D**, respectively, of HBS96. The frontend describes the non-ideal cross-coupling between the receptors, and is typically assumed to remain stable over timescales of the order of days. Therefore, the frontend can be calibrated less frequently than the backend component, which describes the complex gains applied to the receptors. These gains are typically adjusted at the start of each observation, especially if the system equivalent flux density varies strongly with position on the sky. HBS96, B2000, and H2001 describe models that reflect the order in which transformations occur, and that can be decomposed into frontend and backend components.

7.2. *Surjectivity*

When considering only linear transformations of the electric field by a single antenna, as represented using Jones matrices, a surjective model must have 7 independent dof. If the absolute gain is treated as a scalar multiplier, then the matrix component of the transformation must be described by six parameters. In the framework developed by B2000, these 6 dof describe three Lorentz boosts that mix the total intensity with Stokes Q, U, and V and three Euclidean rotations about the Stokes Q, U, and V axes.

The model developed by H2001 includes only 5 of these 6 dof: the differential gain, ΔG (eq. 20) describes the mixing between Stokes I and Q; the symmetric cross-coupling amplitude and phase, ϵ and ϕ (eq. 18) describe the mixing of Stokes I with U and V; the differential phase, ψ (eq. 20) describes a rotation about the Stokes Q axis, which mixes Stokes U and V; and the ellipticity angle, α (eq. 15) describes a rotation about the Stokes U axis, which mixes Stokes Q and V. Missing from this list is the rotation about the Stokes V axis that corresponds to physically rotating the receiver about the line of sight, which mixes Stokes Q and U. As noted in H2001, this rotation can be constrained only through observation of a source with an accurately known position angle. The same conclusion is reached in Appendix B of van Straten (2004).

In HBS96, the product of **G** and **D** has 8 dof (four complex-valued matrix elements: g_p, g_q, d_p, and d_q) because the relative phase between each pair of array elements is important when computing the visibility matrix (Eqn. 51). The absolute phase can be eliminated by replacing **G** with its polar decomposition.

7.3. *Injectivity*

A model that does not injectively map the vector space of its parameters to that of the transformations they describe leads to ambiguous solutions. For example, for circularly-polarized receptors, both the differential phase and the orientation of the receiver about the line of sight correspond to rotations about the Stokes V axis. Therefore, it is necessary to ensure that the rotation axes in a model of the receiver are defined with respect to the (S_1, S_2, S_3) basis and not the (Q, U, V) basis.

Section 5.1 of H2001 describes ambiguity in the ellipticity angle, α, and differential phase, ψ, that arises in the case of linearly-polarized receptors. They assert that two possible solutions, $(\alpha_1, \psi_1) = (0, \psi_0)$ and $(\alpha_2, \psi_2) = (\pi/2, \psi_0 + \pi)$, differ by only a 180° rotation about the Stokes V axis, which negates the unknown values of Stokes Q and U of the calibrator source. Referring to Appendix D.3 and Section 4 of B2000, this ambiguity can be eliminated by restricting the ellipticity angle to the interval $-\pi/4 \leq \alpha \leq \pi/4$.

7.4. *Self-consistency*

Many treatments begin with a description of linear transformations of the electric field, as represented by Jones matrices. In this case, the corresponding Mueller matrices should be pure. However, to simplify the manual expansion of products of these Mueller matrices, some authors introduce small-value approximations. These typically result in an impure Mueller matrix that is inconsistent with the assumed linear response to the electric field.

For example, after assuming that the differential gain ΔG is small, H2001 arrive at equation (20),

$$\underline{\underline{\mathbf{M}}}_A = \begin{pmatrix} 1 & \Delta G/2 & 0 & 0 \\ \Delta G/2 & 1 & 0 & 0 \\ 0 & 0 & \cos\psi & -\sin\psi \\ 0 & 0 & \sin\psi & \cos\psi \end{pmatrix}. \quad (105)$$

In principle, this Mueller matrix can be shown to be impure by computing its associated target coherency matrix and applying the test defined by Equation (C35). In this particular case, it is more instructive and equally valid to show that this Mueller matrix cannot be pure because it can transform a purely polarized state into a nonphysical state that fails to satisfy Equation (23).

Consider the observation of an ideal reference source (Eqn. 90) with Stokes parameters, $\underline{\underline{\boldsymbol{S}}} = I[1, 0, 1, 0]^T$. The transformed Stokes parameters,

$$\underline{\underline{\boldsymbol{S}}}' = \underline{\underline{\mathbf{M}}}_A \underline{\underline{\boldsymbol{S}}} = I[1, \Delta G/2, \cos\psi, \sin\psi]^T, \quad (106)$$

have a degree of polarization,

$$p' = \sqrt{1 + \frac{\Delta G^2}{4}} > 1. \quad (107)$$

As no linear transformation of the electric field vector can convert a valid polarization state into an over-polarized state (Proof 4.3), $\underline{\underline{\mathbf{M}}}_A$ cannot have an equivalent Jones matrix and therefore must be impure. Not all impure Muller matrices result in over-polarization; some impure transformations depolarize a purely polarized state (e.g. see Appendix C.7), and no linear transformation of the electric field is able to do so (Proof 4.4). This specific example demonstrates the potential pitfalls of using Mueller matrices and small-value approximations during numerical analysis.

7.5. *Numerical Stability*

Numerical stability of a mathematical model is critical when fitting the model to experimental data. The fundamental problems caused by numerical instability impact on any method of model fitting; however, for the purpose of concrete illustration, this section and the analysis in Appendix I focus on techniques that optimize a merit function by inverting a Hessian matrix that encodes its local curvature. In this case, a model is unstable when the Hessian matrix becomes ill-conditioned (i.e., prone to large numerical errors during inversion) or singular (i.e., non-invertible). The Hessian is ill-conditioned when two or more model parameters are highly collinear; it is singular when the merit function is effectively independent of one or more of its parameters, such that the partial derivatives of the model with respect to its degenerate parameters are zero.

Three types of numerical instability – fundamental, extrinsic, and intrinsic – are distinguished as follows. Fundamental instability arises when the available data do not constrain all of the dof in a surjective model; without any further assumptions or constraints, the degenerate dof will cause numerical methods to fail. An extrinsically unstable model provides no means of isolating degenerate dof, such that necessary constraining assumptions can be introduced when fundamental instability is encountered. A model is intrinsically unstable if it causes numerical methods to fail even when there are sufficient constraints on all dof.

7.5.1. *Fundamental Instability*

Appendix B of van Straten (2004) describes the fundamental instability caused by 2 degenerate dof that arise when the only available experimental constraints are observations of unknown sources made at multiple parallactic angles. In H2001, these degeneracies are avoided by introducing two assumptions. The rotation about the line of sight is assumed to be zero; and the calibrator source is assumed to have zero circular polarization. The latter assumption is typically invalid when the calibrator source is a pulsar; therefore, additional observations of other sources with either known or assumed circular polarization must be incorporated to constrain the instrumental mixing between I and V (e.g., Liao et al. 2016).

The analysis of fundamental instability is extended in Appendix I, which identifies additional degenerate dof in impure Mueller matrices, owing to unknown components of the instrumental response that commute with the matrix argument. It also presents an example of parameter collinearity that arises when an unknown component of the Stokes parameters intrinsic to the observed sources is an eigenvector of the matrix argument. Owing to the degenerate sign of Stokes V identified in Appendix I.1, the phase convention of the backend cannot be determined using only unknown sources observed at multiple parallactic angles, regardless of the basis defined by the polarization of the receptors.

Section 5.2 of H2001 identifies this degeneracy only for the circular basis, and incorrectly concludes that two solutions, one with $\alpha_1 = -\pi/4$ and the other with $\alpha_2 = \pi/4$, differ only by negation of Stokes Q. Stokes V is also negated by the 180° rotation about the Stokes U axis that describes the transformation from α_1 to α_2. H2001 also incorrectly asserts that negating Stokes Q rotates the position angle by 90°. Negating Stokes Q also negates the position angle, such that $\mathrm{P.A.}_1 = \pi/2 - \mathrm{P.A.}_2$.

7.5.2. *Extrinsic Instability*

Extrinsic instability arises when the degenerate dof in a model cannot be isolated from other dof. For example, the 4 dof of the frontend transformation must effectively model two Euclidean rotations and two Lorentz boosts. When the receptors are linearly polarized, these correspond to a rotation about and a boost along the Stokes U axis, and 2 potentially degenerate dof including a rotation about and a boost along the Stokes V axis. These rotations and boosts are clearly separated in equation (19) of B2000. In H2001, the Stokes V boost is effectively differentiated from the Stokes U boost by the cross-coupling phase, and the Stokes V rotation is assumed to be zero. In the deviations from an ideal feed **D** employed by HBS96, the 4 dof are inseparably combined in the two complex-valued elements, d_p, and d_q. Consequently, the HBS96 model can be applied to single-antenna observations of point sources only if sufficient observational constraints eliminate the known degeneracies.

7.5.3. *Intrinsic Instability*

Appendix II of Conway & Kronberg (1969) describes an intrinsically unstable model of undesirable cross coupling between the two receptors of an imperfect feed. This model is adapted in equation (A2) of Stinebring et al. (1984) and equation (16) of H2001,

$$\boldsymbol{e}' = \begin{pmatrix} 1 & \epsilon_1 e^{i\phi_1} \\ \epsilon_2 e^{-i\phi_2} & 1 \end{pmatrix} \boldsymbol{e}. \tag{108}$$

The equation is intrinsically unstable in the vicinity of the ideal solution owing to the polar coordinate singularity at the origin, where the phase angle ϕ_k becomes undetermined as the cross-coupling amplitude ϵ_k approaches zero. This problem persists in equation (18) of H2001, where ϕ is poorly constrained when ϵ is small. Johnston (2002) addresses this problem by directly modeling the real and imaginary components of all complex-valued quantities, instead of their amplitude and phase. In contrast, equation (15) of B2000 is unstable only at the poles, where $\chi = \pm\pi/4$, and the orientation θ becomes degenerate with the differential phase. When the basis is correctly defined, these poles are well away from the region in which the anticipated solution lies.

8. CONCLUSION

The geometry of the polarization ellipse that defines the Stokes parameters is elegantly connected to that of the coherency matrix via the Pauli matrices. Geometry also relates linear transformations of the electric field to their impact on the Stokes parameters, leading to a physically meaningful and broadly applicable classification of polarimetric transformations. The properties of these transformations simplify the analysis of the elements of the signal path and yield both theoretical insights and practical guidelines that can be applied during calibration.

These guidelines include criteria for evaluating and selecting a model of the instrumental response that performs well during numerical analysis of experimental data. Among those considered, the model introduced by Britton (2000) satisfies all of the selection criteria for single-antenna observations of a point source. In this physically-motivated and self-consistent model, there is a one-to-one mapping between parameters and linear transformations, which are described using equations that remain numerically stable during modeling.

This introduction focuses primarily on the mathematical and conceptual foundations of polarimetry, with emphasis on the numerical analysis of radio astronomical observations. For a broader review of both the physical processes that produce polarized radiation and the astrophysical phenomena that can be studied through polarimetry, the review by Trippe (2014) is exceptionally thorough and insightful.

Radio polarimetry is a vibrant field of research that continues to yield transformative insights and discoveries in astrophysics, driven by both technological advances and innovative techniques. For example, the relatively recent discovery of polarization closure traces (Broderick & Pesce 2020; Samuel et al. 2022) enabled the first images of magnetic field structure near the M87 black hole using very long baseline interferometry (Event Horizon Telescope Collaboration et al. 2021). Novel techniques of high-fidelity polarimetry led to the discovery of a surprisingly high degree of linear polarization in low-frequency radio images of solar bursts (Dey et al. 2025). At even higher time resolution, measurements of the polarized radiation from pulsars, magnetars and fast radio bursts continue to reveal new insights into their nature and extreme environments (e.g., Camilo et al. 2006; Eatough et al. 2013; Michilli et al. 2018; Primak et al. 2022).

Current and future scientific goals, such as detecting the neutral Hydrogen emission from the epoch of reionization (e.g., Trott et al. 2018; Mertens et al. 2020) and understanding the origin and evolution of interstellar and intergalactic magnetic fields (e.g., Han 2017), motivate the development of next-generation facilities and methods of polarimetric analysis, such as Faraday tomography (e.g., Ganguly et al. 2000; Van Eck et al. 2017). Furthermore, the unprecedented sensitivity, resolution, and fields of view of new observatories necessi-

tate further research and development of novel methods of direction-dependent interferometric calibration (e.g., Smirnov & Tasse 2015) and correction of ionospheric Faraday rotation (e.g., de Gasperin et al. 2019). Improved polarimetric fidelity also has the potential to increase the sensitivity of pulsar timing array experiments to the low-frequency stochastic gravitational wave background (e.g., Guillemot et al. 2023; Rogers et al. 2024; Guillemot et al. 2025).

Given the breadth and potential impact of scientific knowledge that remains to be discovered through polarimetry, I hope that this introduction might help others to further pursue the subject.

ACKNOWLEDGMENTS

I'm grateful to the mentors who sparked my interest in polarimetry and the colleagues who share their enthusiasm for the subject, especially Matthew Britton, Paul Demorest, Russell Edwards, Marisa Geyer, Lucas Guillemot, Johan Hamaker, JinLin Han, Simon Johnston, Aris Karastergiou, Vivek Venkatraman Krishnan, Dick Manchester, Maaijke Mevius, Aris Noutsos, Steve Ord, Stefan Osłowski, Aditya Parthasarathy, John Reynolds, Maciej Serylak, Ben Stappers, and Caterina Tiburzi. Marisa's visit in 2018 prompted the development of this tutorial, and Caterina and Maaijke provided insightful feedback and brilliant ideas that significantly enhanced its elaboration. The comprehensive and constructive criticism of the anonymous reviewer led to further additions and improvements.

APPENDIX

A. SIGNAL PROCESSING FUNDAMENTALS

This section reviews the theoretical basis for using complex numbers to represent signals and Jones matrices to represent linear transformations of the electric field.

A.1. *The Analytic Signal*

The voltage signal from each receptor is a real-valued function of time, or process, that may be represented by its associated analytic signal. The analytic signal, also known as Gabor's complex signal, is a complex-valued representation of a real-valued process that provides its instantaneous amplitude and phase. The analytic signal associated with a process, $x(t)$, is defined by

$$z(t) = x(t) + i\hat{x}(t), \tag{A1}$$

where $\hat{x}(t)$ is the Hilbert transform of $x(t)$ (Papoulis 1965). Following a Fourier transform, defined by

$$X(\nu) = \int_{\infty}^{\infty} x(t) \exp(-i2\pi\nu t) dt, \tag{A2}$$

the Hilbert transform is equivalent to

$$\hat{X}(\nu) = H(\nu)X(\nu), \tag{A3}$$

where

$$H(\nu) = \begin{cases} -i & \nu > 0 \\ i & \nu < 0 \end{cases} \tag{A4}$$

is known as the quadrature filter (Papoulis 1965). The quadrature filter can be understood as a 90° phase shifter by noting that the Hilbert transform of $\cos(\nu_0 t)$ is equal to $\sin(\nu_0 t)$ (see Table 2). Using the quadrature filter, it can also be shown that the Fourier transform of the analytic signal, $Z(\nu)$, is equal to zero for $\nu < 0$.

$$\begin{aligned} Z(\nu) &= X(\nu) + i\hat{X}(\nu) \\ &= X(\nu) + iH(\nu)X(\nu) \\ &= \begin{cases} 2X(\nu) & \nu > 0 \\ 0 & \nu < 0 \end{cases} \end{aligned} \tag{A5}$$

Conversely, the analytic signal associated with $x(t)$ is produced by suppressing the negative frequencies in $X(\nu)$.

As it is derived from the real-valued process, the analytic signal does not contain any additional information. However, the analytic signals associated with two orthogonal senses of polarization, $e_0(t)$ and $e_1(t)$, facilitate calculation of the coherency matrix.

A.2. *Narrow-band Approximation*

Presented with a scalar input signal, $e(t)$, the output of a one-dimensional linear time-invariant system is given by the convolution,

$$e'(t) = (j * e)(t) \equiv \int_{-\infty}^{\infty} x(\tau) j(t-\tau) d\tau. \tag{A6}$$

where $j(t)$ is the one-dimensional impulse response function. In a two-dimensional system, each output signal is given by a linear combination of the input signals,

$$\begin{aligned} e'_1(t) &= (j_{00} * e_0)(t) + (j_{01} * e_1)(t), \\ e'_2(t) &= (j_{10} * e_0)(t) + (j_{11} * e_1)(t). \end{aligned} \tag{A7}$$

Table 2. Useful Fourier Transform pairs. The left column lists functions of time. In the right column, the corresponding Fourier transform is given as a function of oscillation frequency, ν.

$x(t)$	$X(\nu)$
$\cos(2\pi\nu_0 t)$	$(\delta(\nu+\nu_0)+\delta(\nu-\nu_0))/2$
$\sin(2\pi\nu_0 t)$	$i(\delta(\nu+\nu_0)-\delta(\nu-\nu_0))/2$
Quadrature Filter	
$h(t)=(\pi t)^{-1}$	$H(\nu)=\begin{cases}-i & \nu>0\\ i & \nu<0\end{cases}$
Lowpass Filter	
$\pi(t)=\mathrm{sinc}(\pi\Delta\nu t)$	$\Pi(\nu/\Delta\nu)=\begin{cases}0 & \lvert\nu/\Delta\nu\rvert>1/2\\ 1/2 & \lvert\nu/\Delta\nu\rvert=1/2\\ 1 & \lvert\nu/\Delta\nu\rvert<1/2\end{cases}$
Shift Theorem	
$x'(t)=x(t+\tau)$	$X'(\nu)=X(\nu)\exp(i\pi\nu\tau)$

Let $e_0(t)$ and $e_1(t)$ be the components of the electric field vector $\boldsymbol{e}(t)$ and $j_{mn}(t)$ be the elements of the 2×2 impulse response matrix, $\mathbf{j}(t)$. By the convolution theorem, Equation (A7) is equivalent to

$$\boldsymbol{E}'(\nu)=\mathbf{J}(\nu)\boldsymbol{E}(\nu), \tag{A8}$$

where $\mathbf{J}(\nu)$ and $\boldsymbol{E}(\nu)$ are the Fourier transforms of $\mathbf{j}(t)$ and $\boldsymbol{e}(t)$. Under the assumption that $\mathbf{J}(\nu)$ is approximately constant over a sufficiently narrow range of frequencies, convolution reduces to multiplication by a Jones matrix.

B. VECTOR, MATRIX, AND TENSOR NOTATION

The following typographical conventions are used to indicate the dimensions of mathematical symbols. Scalar quantities such as the real value x or the complex value z are typeset in italics. All vectors are typeset with a bold italic font. Two-dimensional vectors, such as the electric field $\boldsymbol{e}$, are not underlined. Three-dimensional vectors, such as the Stokes polarization vector $\underline{\boldsymbol{S}}$, have a single underline. Four-dimensional vectors, such as the Stokes four-vector $\underline{\underline{\boldsymbol{S}}}$, have a double underline.

All matrices have a bold roman font; 2×2 matrices like the Jones matrix $\mathbf{J}$ are not underlined; 3×3 matrices like the depolarizer matrix $\underline{\mathbf{M}}_\Delta$ are underlined once; and 4×4 matrices like the Mueller matrix $\underline{\underline{\mathbf{M}}}$ are underlined twice. Rank 4 tensors, typeset using a bold calligraphic font, include the two-dimensional polarization transfer tensor $\boldsymbol{\mathcal{U}}$ and the four-dimensional $\underline{\underline{\boldsymbol{\mathcal{T}}}}$ in Appendix C.6.

Unit vectors, such as a receptor with unity gain $\hat{\boldsymbol{r}}$ or a rotation axis $\hat{\underline{\boldsymbol{n}}}$, are decorated with a hat. A tilde is used to denote both singular matrices and null four-vectors, such as the instantaneous coherency matrix $\tilde{\boldsymbol{\rho}}$ and Stokes parameters $\tilde{\underline{\underline{\boldsymbol{s}}}}$ computed from a single instance of the electric field.

The elements of a column vector are labeled with a sub-script, as in the two components of the electric field,

$$\boldsymbol{e}=\begin{pmatrix}e_x\\ e_y\end{pmatrix}. \tag{B1}$$

and the four components of the Stokes parameters, S_μ. To distinguish between Jones vectors that represent the electric field and those that represent receptors, the latter are described using row vectors with a super-scripted index; e.g.

$$\boldsymbol{r}=(r^x,\,r^y). \tag{B2}$$

The Hermitian transpose converts a column vector to a row vector (and vice versa) and converts each component to its complex conjugate; e.g.

$$\boldsymbol{e}^\dagger=\left(e_x^*,\,e_y^*\right). \tag{B3}$$

The scalar product (or dot product) between a row vector and a column vector is given by

$$\boldsymbol{r}\cdot\boldsymbol{e}=r^x e_x+r^y e_y=r^i e_i, \tag{B4}$$

where the last equality uses the Einstein convention that implies summation over the repeated index i. A scalar product is implied when a row vector on the left is multiplied by a column vector on the right; i.e.,

$$\boldsymbol{r}\boldsymbol{e}=\boldsymbol{r}\cdot\boldsymbol{e}=r^i e_i. \tag{B5}$$

Using the same index notation, the element in the $\mu^{\rm th}$ row and $\nu^{\rm th}$ column of a matrix $\mathbf{A}$ is

$$A^\nu_\mu \tag{B6}$$

and the $\mu^{\rm th}$ row and $\nu^{\rm th}$ column of a matrix product is given by the sum

$$\{\mathbf{A}\cdot\mathbf{B}\}^\nu_\mu=A^\gamma_\mu B^\nu_\gamma. \tag{B7}$$

Similarly, the $\mu^{\rm th}$ row of a matrix times a column vector is given by

$$\{\mathbf{A}\boldsymbol{x}\}_\mu=A^\gamma_\mu x_\gamma. \tag{B8}$$

Vectors and matrices can be seen as rank 1 and rank 2 tensors, respectively, where the rank of a tensor is equal to the number of indeces required to define its elements. The elements of a rank 4 tensor are described using four indeces; e.g.

$$A^{\nu\lambda}_{\mu\kappa}. \tag{B9}$$

Given two tensors, a rank N tensor A and a rank M tensor B, the tensor product $A \otimes B$ yields a rank $N+M$ tensor. For example, the tensor product of two matrices (each a rank 2 tensor) yields a rank 4 tensor with elements given by

$$\{\mathbf{A} \otimes \mathbf{B}\}^{\nu\lambda}_{\mu\kappa} = A^{\nu}_{\mu} B^{\lambda}_{\kappa}. \tag{B10}$$

The tensor product of a column vector and a row vector (each a rank 1 tensor) yields a matrix (rank 2 tensor) with elements

$$\{\boldsymbol{e} \otimes \boldsymbol{r}\}^{j}_{i} = e_i r^j. \tag{B11}$$

A tensor product is implied when a column vector on the left is multiplied by a row vector on the right; e.g.

$$\boldsymbol{e}\boldsymbol{r} = \boldsymbol{e} \otimes \boldsymbol{r}. \tag{B12}$$

A contraction of two tensors, $A \cdot B$, is defined as a tensor product followed by summation over a pair of matching indeces, thereby reducing the rank of the result of the tensor product by two. For example, the scalar product of a row vector and column vector is equivalent to a contraction. Each is a rank 1 tensor and their tensor product yields a rank 2 tensor; summation over their indeces, as in Equation (B4), yields a scalar (rank 0 tensor). Similarly, the product of two matrices is equivalent to a tensor contraction. Each is a rank 2 tensor and their tensor product yields a rank 4 tensor; summation over a pair of indeces, as in Equation (B7), yields a rank 2 tensor (another matrix).

A double contraction of two tensors, $A{:}B$, is defined as a tensor product followed by summation over two pairs of matching indeces, thereby reducing the rank of the result of the tensor product by four. For example, the double contraction of a rank 4 tensor $\boldsymbol{\mathcal{U}}$ with a (rank 2) matrix $\mathbf{C}$ yields another (rank 2) matrix with elements given by

$$\{\mathcal{U} : \mathbf{C}\}^{\nu}_{\mu} = \mathcal{U}^{\nu\lambda}_{\mu\kappa} C^{\kappa}_{\lambda}. \tag{B13}$$

Using the Einstein summation convention, the trace of a matrix $\mathbf{A}$ is

$$\mathrm{Tr}\,[\mathbf{A}] = A^{\beta}_{\beta} \tag{B14}$$

and the trace of a matrix product

$$\mathrm{Tr}\,[\mathbf{AB}] = \mathrm{Tr}\,[\mathbf{BA}] = A^{\gamma}_{\beta} B^{\beta}_{\gamma} = \mathbf{A} : \mathbf{B} = \mathbf{B} : \mathbf{A}. \tag{B15}$$

That is, the trace of a matrix product is equivalent to the double contraction of the matrices, also known as the projection with respect to the trace inner product. Like the scalar product of two vectors, this projection is symmetric and matrices $\mathbf{A}$ and $\mathbf{B}$ are orthogonal with respect to the trace inner product if $\mathbf{A} : \mathbf{B} = 0$.

Equations (46) and (47) introduce transformation properties exhibited by two tensor double contractions. Equation (46) follows from the definition of the tensor product (Eqns. [44] and [B10]).

Proof B.1.

$$\begin{aligned}
\{(\mathbf{A} \otimes \mathbf{B}) : \mathbf{C}\}^{\nu}_{\mu} &= \{\mathbf{A} \otimes \mathbf{B}\}^{\nu\lambda}_{\mu\kappa} C^{\kappa}_{\lambda} && \text{Eqn. (B13)} \\
&= A^{\nu}_{\mu} B^{\lambda}_{\kappa} C^{\kappa}_{\lambda} && \text{Eqn. (B10)} \\
&= A^{\nu}_{\mu} (\mathbf{B} : \mathbf{C}) && \text{Eqn. (B15)} \\
&= \{\mathbf{A}\,(\mathbf{B} : \mathbf{C})\}^{\nu}_{\mu}
\end{aligned}$$

Likewise, Equation (47) follows from the definition of the $\tilde{\otimes}$ operator (Eqn. 45).

Proof B.2.

$$\begin{aligned}
\left\{(\mathbf{A} \,\tilde{\otimes}\, \mathbf{B}) : \mathbf{C}\right\}^{\nu}_{\mu} &= \left\{\mathbf{A} \,\tilde{\otimes}\, \mathbf{B}\right\}^{\nu\lambda}_{\mu\kappa} C^{\kappa}_{\lambda} && \text{Eqn. (B13)} \\
&= A^{\lambda}_{\mu} B^{\nu}_{\kappa} C^{\kappa}_{\lambda} && \text{Eqn. (45)} \\
&= A^{\lambda}_{\mu} \{\mathbf{C} \cdot \mathbf{B}\}^{\nu}_{\lambda} && \text{Eqn. (B7)} \\
&= \{\mathbf{ACB}\}^{\nu}_{\mu} && \text{Eqn. (B7)}
\end{aligned}$$

C. LINEAR ALGEBRA FUNDAMENTALS

This section reviews the elements of linear algebra used in this work, beginning with the Hermitian basis matrices. The spectral decomposition is used to derive the axis-angle representations of rotations and boosts, the symmetry properties of which are used to identify matrices that commute and the eigenmatrices of a congruence transformation. Finally, the symmetry between Equations (44) and (45) is used to define a pure Mueller matrix and its associated Jones matrix, and an example of a depolarizing impure Mueller matrix is presented.

C.1. *Hermitian Basis Matrices*

The Hermitian basis matrices (Eqn. 15) have a number of useful algebraic properties that enable meaningful geometric interpretations of the equations used in this work. When the basis matrices are indexed with a Greek character, the index is understood to span all dimensions including the identity; e.g., $\mu \in \{0, 1, 2, 3\}$. When Latin characters are used, the index is understood to span only the Pauli matrces; e.g., $m \in \{1, 2, 3\}$. The

following identities prove useful in this work.

$$\mathrm{Tr}\,[\boldsymbol{\sigma}_i] = 0 \tag{C1}$$
$$|\boldsymbol{\sigma}_i| = -1 \tag{C2}$$
$$|\boldsymbol{\sigma}_\mu + \boldsymbol{\sigma}_\nu| = |\boldsymbol{\sigma}_\mu| + |\boldsymbol{\sigma}_\nu|; \quad \mu \neq \nu \tag{C3}$$
$$\boldsymbol{\sigma}_\mu : \boldsymbol{\sigma}_\nu = 2\delta_{\mu\nu} \tag{C4}$$
$$\boldsymbol{\sigma}_i\, \boldsymbol{\sigma}_j = \delta_{ij}\, \boldsymbol{\sigma}_0 + i\epsilon_{ijk}\, \boldsymbol{\sigma}_k \tag{C5}$$

In the last identity, ϵ_{ijk} is the permutation symbol and summation over the index k is implied. Equation (C1) expresses the traceless property of the Pauli matrices. Equations (C2) and (C3) can be used to show that

$$|x_\mu \boldsymbol{\sigma}_\mu| = x_\mu^2 |\boldsymbol{\sigma}_\mu| = x_0^2 - x_1^2 - x_2^2 - x_3^2. \tag{C6}$$

That is, the determinant of a linear combination of the Hermitian basis matrices is equal to the invariant interval of the four-vector of the scalar coefficients, x_μ. Equation (C4) states that the Hermitian basis matrices are mutually orthogonal with respect to the trace inner product. The first term on the right-hand side of Equation (C5) identifies the Pauli matrices as the square roots of $\boldsymbol{\sigma}_0$. The second term describes how the Pauli matrices behave like the basis vectors of a right-handed coordinate system, where $\underline{\hat{\boldsymbol{v}}}_i \times \underline{\hat{\boldsymbol{v}}}_j = \epsilon_{ijk}\, \underline{\hat{\boldsymbol{v}}}_k$. Using this identity, the product of two matrices,

$$\begin{aligned} \mathbf{AB} &= (a\,\boldsymbol{\sigma}_0 + \underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{\sigma}})(b\,\boldsymbol{\sigma}_0 + \underline{\boldsymbol{b}} \cdot \underline{\boldsymbol{\sigma}}) \\ &= (ab + \underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{b}})\boldsymbol{\sigma}_0 + (a\underline{\boldsymbol{b}} + b\underline{\boldsymbol{a}} + i\underline{\boldsymbol{a}} \times \underline{\boldsymbol{b}}) \cdot \underline{\boldsymbol{\sigma}}, \end{aligned} \tag{C7}$$

where $\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3)$ is the Pauli vector that appears in Equations (59) and (60).

C.2. *Basis Transformations*

A pair of orthonormal receptors, $\hat{\boldsymbol{r}}_0$ and $\hat{\boldsymbol{r}}_1$, have unit length and are orthogonal; that is, they satisfy

$$\hat{\boldsymbol{r}}_i \cdot \hat{\boldsymbol{r}}_j^\dagger = \delta_{i,j}. \tag{C8}$$

The projection of the electric field vector onto a pair of orthonormal receptors defines a basis transformation that is equivalent to a unitary Jones matrix.

Proof C.1. Let

$$\mathbf{R} = \begin{pmatrix} \hat{\boldsymbol{r}}_0 \\ \hat{\boldsymbol{r}}_1 \end{pmatrix}$$

such that

$$\begin{aligned} \left\{\mathbf{R}\mathbf{R}^\dagger\right\}_i^j &= \{\mathbf{R}\}_i^k \left\{\mathbf{R}^\dagger\right\}_k^j && \text{Eqn. (B7)} \\ &= \{\hat{\boldsymbol{r}}_i\}^k \left\{\hat{\boldsymbol{r}}_j^\dagger\right\}_k \\ &= \hat{\boldsymbol{r}}_i \cdot \hat{\boldsymbol{r}}_j^\dagger && \text{Eqn. (B4)} \\ &= \delta_{i,j} && \text{Eqn. (C8)} \end{aligned}$$

That is, $\mathbf{R}\mathbf{R}^\dagger = \boldsymbol{\sigma}_0$; therefore, $\mathbf{R}^\dagger = \mathbf{R}^{-1}$.

In general, a unitary matrix has a determinant equal to $\exp(i\phi)$, and multiplication by $\exp(-i\phi/2)$ yields a unimodular matrix (with $|\mathbf{R}| = 1$). Basis transformations defined by unitary matrices are used extensively throughout both Section 5 and the following sections of the appendix.

C.3. *Spectral Decomposition*

The eigenvalues λ and eigenvectors $\boldsymbol{v}$ of a matrix $\mathbf{A}$ satisfy the following (logically-equivalent) equations,

$$\mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v} \quad \Leftrightarrow \quad (\mathbf{A} - \lambda \mathbf{I})\, \boldsymbol{v} = 0, \tag{C9}$$

where $\mathbf{I}$ is the identity matrix with the dimensions of $\mathbf{A}$. Equation (C9) can be expressed simultaneously for all eigenvalue/eigenvector pairs by $\mathbf{AR} = \mathbf{R}\boldsymbol{\Lambda}$, where the k^{th} column of $\mathbf{R}$ is the unit eigenvector $\hat{\boldsymbol{v}}_k$, and the diagonal matrix $\boldsymbol{\Lambda}$ is defined by $\Lambda_k^j = \delta_{j,k}\lambda_k$. This can be rearranged to yield the eigendecomposition of $\mathbf{A}$,

$$\mathbf{A} = \mathbf{R}\boldsymbol{\Lambda}\mathbf{R}^{-1}. \tag{C10}$$

If $\mathbf{A}$ is normal (i.e., $\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}\mathbf{A}^\dagger$), and $\boldsymbol{v}$ is an eigenvector of $\mathbf{A}$ with associated eigenvalue λ, then $\boldsymbol{v}$ is also an eigenvector of $\mathbf{A}^\dagger$ with associated eigenvalue λ^*.

Proof C.2. Define $\mathbf{L} = \mathbf{A} - \lambda\mathbf{I}$ and consider

$$\begin{aligned} |\mathbf{L}^\dagger \boldsymbol{v}|^2 &= \left(\mathbf{L}^\dagger \boldsymbol{v}\right)^\dagger \left(\mathbf{L}^\dagger \boldsymbol{v}\right) \\ &= \boldsymbol{v}^\dagger \mathbf{L}\mathbf{L}^\dagger \boldsymbol{v} \\ &= \boldsymbol{v}^\dagger \mathbf{L}^\dagger \mathbf{L} \boldsymbol{v} && \mathbf{L} \text{ is normal} \\ &= (\mathbf{L}\boldsymbol{v})^\dagger\, \mathbf{L}\boldsymbol{v} \\ &= |\mathbf{L}\boldsymbol{v}|^2 \\ &= |\,(\mathbf{A} - \lambda\mathbf{I})\, \boldsymbol{v}|^2 = 0 && \text{Eqn. (C9)} \end{aligned}$$

Therefore $\mathbf{L}^\dagger \boldsymbol{v} = \left(\mathbf{A}^\dagger - \lambda^* \mathbf{I}\right) \boldsymbol{v} = 0$.

Furthermore, if $\mathbf{A}$ is normal, then Equation (C10) is equivalent to a congruence transformation by a unitary matrix known as its spectral decomposition,

$$\mathbf{A} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^\dagger. \tag{C11}$$

Proof C.3. If $\mathbf{A}$ is normal, then

$$\begin{aligned}
(\mathbf{A}^\dagger \hat{\boldsymbol{v}}_i)^\dagger \hat{\boldsymbol{v}}_j &= \hat{\boldsymbol{v}}_i^\dagger \mathbf{A} \hat{\boldsymbol{v}}_j \\
(\lambda_i^* \hat{\boldsymbol{v}}_i)^\dagger \cdot \hat{\boldsymbol{v}}_j &= \hat{\boldsymbol{v}}_i^\dagger \cdot (\lambda_j \hat{\boldsymbol{v}}_j) \qquad \text{Proof C.2} \\
\lambda_i \hat{\boldsymbol{v}}_i^\dagger \cdot \hat{\boldsymbol{v}}_j &= \lambda_j \hat{\boldsymbol{v}}_i^\dagger \cdot \hat{\boldsymbol{v}}_j \\
(\lambda_i - \lambda_j) \hat{\boldsymbol{v}}_i^\dagger \cdot \hat{\boldsymbol{v}}_j &= 0
\end{aligned}$$

If $\lambda_i \neq \lambda_j$, then $\hat{\boldsymbol{v}}_i^\dagger \cdot \hat{\boldsymbol{v}}_j = 0$ and

$$\begin{aligned}
\left\{\mathbf{R}^\dagger \mathbf{R}\right\}_i^j &= \left\{\mathbf{R}^\dagger\right\}_i^k \{\mathbf{R}\}_k^j \\
&= \left\{\hat{\boldsymbol{v}}_i^\dagger\right\}^k \{\hat{\boldsymbol{v}}_j\}_k = \hat{\boldsymbol{v}}_i^\dagger \cdot \hat{\boldsymbol{v}}_j = \delta_{i,j}
\end{aligned}$$

That is, $\mathbf{R}^\dagger \mathbf{R} = \mathbf{I}$; therefore, $\mathbf{R}^\dagger = \mathbf{R}^{-1}$.

In the natural basis defined by $\mathbf{R}^\dagger$, the matrix $\mathbf{A}$ becomes the diagonal matrix,

$$\mathbf{\Lambda} = \mathbf{R}^\dagger \mathbf{A} \mathbf{R}. \tag{C12}$$

Both Hermitian and unitary matrices are normal, and each takes diagonal form in the natural basis defined by its eigenvectors. Consequently, any Hermitian matrix can be diagonalized such that it is equivalent to a differential gain transformation (see Section 5.2.1) in its natural basis. Similarly, any unitary matrix is equivalent to differential phase (see Section 5.2.2) in its natural basis. This is applied in Proof 5.2, where rotation about the line of sight is shown to be equivalent to differential phase between the circularly-polarized natural modes of a magnetized cold plasma.

The spectral decomposition can also be written as a linear combination of outer products,

$$\mathbf{A} = \lambda_k \, \hat{\boldsymbol{v}}_k \otimes \hat{\boldsymbol{v}}_k^\dagger. \tag{C13}$$

Proof C.4.

$$\begin{aligned}
A_i^j &= \left\{\mathbf{R}\mathbf{\Lambda}\mathbf{R}^\dagger\right\}_i^j && \text{Eqn. (C11)} \\
&= \{\mathbf{R}\}_i^k \{\mathbf{\Lambda}\}_k^l \left\{\mathbf{R}^\dagger\right\}_l^j && \text{Eqn. (B7)} \\
&= \{\hat{\boldsymbol{e}}_k\}_i \; \delta_{k,l} \; \lambda_k \left\{\hat{\boldsymbol{e}}_l^\dagger\right\}^j \\
&= \lambda_k \; \{\hat{\boldsymbol{e}}_k\}_i \left\{\hat{\boldsymbol{e}}_k^\dagger\right\}^j \\
&= \lambda_k \left\{\hat{\boldsymbol{e}}_k \otimes \hat{\boldsymbol{e}}_k^\dagger\right\}_i^j && \text{Eqn. (B12)}
\end{aligned}$$

If the eigenvalues are distinct, then

$$\mathbf{P}_k = \hat{\boldsymbol{v}}_k \otimes \hat{\boldsymbol{v}}_k^\dagger \tag{C14}$$

form a complete set of mutually orthogonal projection matrices; i.e.,

$$\mathbf{P}_i \mathbf{P}_j = \delta_{i,j} \mathbf{P}_j \tag{C15}$$

and

$$\sum_i \mathbf{P}_i = \mathbf{I} \tag{C16}$$

Proof C.5. $\mathbf{P}_k$ are orthogonal and idempotent.

$$\begin{aligned}
\mathbf{P}_i \mathbf{P}_j &= (\hat{\boldsymbol{v}}_i \hat{\boldsymbol{v}}_i^\dagger)(\hat{\boldsymbol{v}}_j \hat{\boldsymbol{v}}_j^\dagger) && \text{Eqn. (B12)} \\
&= \hat{\boldsymbol{v}}_i (\hat{\boldsymbol{v}}_i^\dagger \hat{\boldsymbol{v}}_j) \hat{\boldsymbol{v}}_j^\dagger && \text{Associativity} \\
&= \delta_{i,j} \hat{\boldsymbol{v}}_i \hat{\boldsymbol{v}}_j^\dagger && \hat{\boldsymbol{v}}_i^\dagger \hat{\boldsymbol{v}}_j = \delta_{i,j} \\
&= \delta_{i,j} \mathbf{P}_j && \text{Eqn. (C14)}
\end{aligned}$$

Proof C.6. $\mathbf{P}_k$ form a complete set.
Let $\boldsymbol{w} = c_j \hat{\boldsymbol{v}}_j$ and consider

$$\begin{aligned}
\sum_i \mathbf{P}_i \boldsymbol{w} &= \left(\hat{\boldsymbol{v}}_i \hat{\boldsymbol{v}}_i^\dagger\right)(c_j \hat{\boldsymbol{v}}_j) && \text{Eqn. (C14)} \\
&= c_j \hat{\boldsymbol{v}}_i \hat{\boldsymbol{v}}_i^\dagger \hat{\boldsymbol{v}}_j && \text{Commutativity} \\
&= c_j \hat{\boldsymbol{v}}_i \delta_{i,j} && \hat{\boldsymbol{v}}_i^\dagger \hat{\boldsymbol{v}}_j = \delta_{i,j} \\
&= c_j \hat{\boldsymbol{v}}_j \\
&= \boldsymbol{w}
\end{aligned}$$

The 2×2 Hermitian coherency matrix has two eigenvalues, $\lambda_k = (S_0 \pm |\underline{\boldsymbol{S}}|)/2$, and can be expressed as a spectral decomposition,

$$\begin{aligned}
\boldsymbol{\rho} &= \lambda_0 \, \hat{\boldsymbol{e}}_0 \otimes \hat{\boldsymbol{e}}_0^\dagger + \lambda_1 \, \hat{\boldsymbol{e}}_1 \otimes \hat{\boldsymbol{e}}_1^\dagger \\
&= \lambda_0 \, \tilde{\boldsymbol{\rho}}_0 + \lambda_1 \, \tilde{\boldsymbol{\rho}}_1,
\end{aligned} \tag{C17}$$

where $\tilde{\boldsymbol{\rho}}_k = \hat{\boldsymbol{e}}_k \otimes \hat{\boldsymbol{e}}_k^\dagger$ correspond to purely polarized states, as in Equation (24). That is, any partially polarized state can be represented as an incoherent superposition of a pair of purely polarized states that are

orthogonally polarized ($\hat{e}_0^\dagger \hat{e}_1 = 0$ and $\tilde{\boldsymbol{\rho}}_0 \tilde{\boldsymbol{\rho}}_1 = 0$). If the signal is unpolarized, then the eigenvalues are equal and any non-zero vector is an eigenvector. If $\boldsymbol{\rho}$ is purely polarized, then one of the eigenvalues equals zero, and the polarization state is completely described by the Jones vector associated with the non-zero eigenvalue.

In the natural basis defined by $\mathbf{R}^\dagger$, the eigenvalues λ_m are equal to the variances of two uncorrelated signals received by orthogonally polarized receptors described by the Hermitian transposes of the eigenvectors. In this basis, the total intensity, $S_0 = \lambda_0 + \lambda_1$; the polarized intensity, $S_1 = |\underline{\boldsymbol{S}}| = \lambda_0 - \lambda_1$; and $S_2 = S_3 = 0$. That is, $\mathbf{R}^\dagger$ rotates the basis such that the polarization vector points along S_1. In this basis, it is clear that the degree of polarization

$$p = \frac{|\underline{\boldsymbol{S}}|}{S_0} = \frac{|\lambda_0 - \lambda_1|}{\lambda_0 + \lambda_1}, \tag{C18}$$

is equal to zero when the intensities of the orthogonal modes are equal ($\lambda_0 = \lambda_1$); in contrast, $p = 1$ when the signal consists of a single purely polarized mode (one of the eigenvalues is zero).

C.4. *Axis-angle Representation*

Equations (59) and (60) are defined using the matrix exponential, which is given by the power series,

$$\exp(\mathbf{A}) \equiv \sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k!}. \tag{C19}$$

Here, $\mathbf{A}^k$ is the k^{th} power of $\mathbf{A}$; $\mathbf{A}^0 = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix with the dimensions of $\mathbf{A}$; and $k!$ is the factorial of k. If $\mathbf{A}$ is normal, then its spectral decomposition (Eqns. [C13] and [C14]),

$$\mathbf{A} = \lambda_k \mathbf{P}_k, \tag{C20}$$

such that

$$\mathbf{A}^n = (\lambda_k)^n \mathbf{P}_k. \tag{C21}$$

Proof C.7. Using induction, start with the base case and apply Equation (C16) to yield

$$\mathbf{A}^0 = (\lambda_k)^0 \mathbf{P}_k = \sum_k \mathbf{P}_k = \mathbf{I}$$

If $\mathbf{A}^n = (\lambda_k)^n \mathbf{P}_k$, then

$$\begin{aligned} \mathbf{A}^{n+1} &= \mathbf{A}\mathbf{A}^n \\ &= (\lambda_j \mathbf{P}_j)\left((\lambda_k)^n \mathbf{P}_k\right) \\ &= \lambda_j (\lambda_k)^n \mathbf{P}_j \mathbf{P}_k \\ &= \lambda_j (\lambda_k)^n \delta_{jk} \mathbf{P}_k \qquad \text{Eqn. (C15)} \\ &= (\lambda_k)^{n+1} \mathbf{P}_k. \end{aligned}$$ □

Substitute Equation (C21) into Equation (C19) to yield

$$\exp(\mathbf{A}) = \left(\sum_{k=1}^{\infty} \frac{(\lambda_j)^n}{k!} \right) \mathbf{P}_j = \exp(\lambda_j) \mathbf{P}_j. \tag{C22}$$

To arrive at Equations (59) and (60), consider the eigenvalues of $\mathbf{A} = \underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{\sigma}}$, which must satisfy

$$|\mathbf{A} - \lambda \mathbf{I}| = |\underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{\sigma}} - \lambda \boldsymbol{\sigma}_0| = \lambda^2 - |\underline{\boldsymbol{a}}|^2 = 0. \tag{C23}$$

That is, there are two eigenvalues given by $\lambda = \pm |\underline{\boldsymbol{a}}|$. Let $a = |\underline{\boldsymbol{a}}|$ and $\hat{\underline{\boldsymbol{a}}} = \underline{\boldsymbol{a}}/a$, and use the spectral decomposition,

$$\mathbf{A} = a\mathbf{P}_0 - a\mathbf{P}_1, \tag{C24}$$

to show that

$$\mathbf{P}_0 + \mathbf{P}_1 = \boldsymbol{\sigma}_0 \quad \text{and} \quad \mathbf{P}_0 - \mathbf{P}_1 = \hat{\underline{\boldsymbol{a}}} \cdot \underline{\boldsymbol{\sigma}}; \tag{C25}$$

therefore

$$\begin{aligned} e^{\underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{\sigma}}} &= \exp(a)\mathbf{P}_0 + \exp(-a)\mathbf{P}_1 \\ &= \cosh(a)(\mathbf{P}_0 + \mathbf{P}_1) + \sinh(a)(\mathbf{P}_0 - \mathbf{P}_1) \\ &= \cosh(a)\boldsymbol{\sigma}_0 + \sinh(a)\hat{\underline{\boldsymbol{a}}} \cdot \underline{\boldsymbol{\sigma}}. \end{aligned} \tag{C26}$$

Setting $\underline{\boldsymbol{a}} = \beta \underline{\hat{\boldsymbol{m}}}$ in Equation (C26) yields the axis-angle representation of Hermitian matrices (Eqn. 59); likewise, setting $\underline{\boldsymbol{a}} = i\phi \underline{\hat{\boldsymbol{n}}}$ yields the axis-angle representation of unitary matrices (Eqn. 60).

C.5. *Congruence Eigenmatrices*

The unit vector in the axis-angle representation of a matrix defines an axis of symmetry in the three-dimensional space of the Stokes polarization vector. This symmetry axis can be exploited to rearrange and simplify matrix equations and identify degenerate systems of equations (as in Appendix B of van Straten 2004, and Appendix I.1 of this paper). For example, two matrices commute when their symmetry axes are collinear (parallel or anti-parallel).

Proof C.8. Given

$$\mathbf{A} = a\,\boldsymbol{\sigma}_0 + \underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{\sigma}} \quad \text{and} \quad \mathbf{B} = b\,\boldsymbol{\sigma}_0 + \underline{\boldsymbol{b}} \cdot \underline{\boldsymbol{\sigma}},$$

Equation (C7) shows that $\mathbf{AB} = \mathbf{BA}$ if and only if $\underline{\boldsymbol{a}} \times \underline{\boldsymbol{b}} = 0$.

Furthermore, if $\mathbf{A}$ and $\mathbf{B}$ commute, then they have common eigenvectors.

Proof C.9. If $\mathbf{AB} = \mathbf{BA}$ and $\mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v}$, then

$$\mathbf{AB}\boldsymbol{v} = \mathbf{BA}\boldsymbol{v} = \lambda \mathbf{B}\boldsymbol{v}.$$

That is, $\mathbf{B}\boldsymbol{v}$ is an eigenvector of $\mathbf{A}$ with the same eigenvalue λ; therefore, $\mathbf{B}\boldsymbol{v} = \mu \boldsymbol{v}$.

The coherency matrix has an axis of symmetry defined by its associated Stokes polarization vector, which facilitates the identification of the eigenmatrices of a congruence transformation. A matrix $\boldsymbol{\rho}$ is a congruence eigenmatrix of an invertible matrix $\mathbf{J}$ with real-valued congruence eigenvalue κ if it satisfies the relation

$$\mathbf{J}\boldsymbol{\rho}\mathbf{J}^\dagger = \kappa\boldsymbol{\rho}. \tag{C27}$$

The eigenvectors of a Jones matrix define the singular congruence eigenmatrices of that Jones matrix.

Proof C.10. If $\mathbf{J}\boldsymbol{e} = \lambda\boldsymbol{e}$, where $\lambda \in \mathbb{C}$, then

$$\begin{aligned}
\mathbf{J}\tilde{\boldsymbol{\rho}}\mathbf{J}^\dagger &= \mathbf{J}\boldsymbol{e}\otimes\boldsymbol{e}^\dagger\mathbf{J}^\dagger && \text{Eqn. (24)}\\
&= \mathbf{J}\boldsymbol{e}\otimes(\mathbf{J}\boldsymbol{e})^\dagger && (\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger\\
&= \lambda\,\boldsymbol{e}\otimes\lambda^*\boldsymbol{e}^\dagger &&\\
&= |\lambda|^2\tilde{\boldsymbol{\rho}} && \tilde{\boldsymbol{\rho}} \equiv \boldsymbol{e}\otimes\boldsymbol{e}^\dagger\\
&= \kappa\tilde{\boldsymbol{\rho}} && \kappa \equiv |\lambda|^2 \in \mathbb{R}
\end{aligned}$$

Accordingly, a Jones matrix has two congruence eigenmatrices.[7] Conversely, if a singular coherency matrix $\hat{\boldsymbol{\rho}} = \boldsymbol{e}\otimes\boldsymbol{e}^\dagger$ is a congruence eigenmatrix of $\mathbf{J}$, then $\boldsymbol{e}$ is an eigenvector of $\mathbf{J}$.

Proof C.11. If $\hat{\boldsymbol{\rho}} = \boldsymbol{e}\otimes\boldsymbol{e}^\dagger$ and $\mathbf{J}\hat{\boldsymbol{\rho}}\mathbf{J}^\dagger = \kappa\hat{\boldsymbol{\rho}}$, then

$$\mathbf{J}\left(\boldsymbol{e}\otimes\boldsymbol{e}^\dagger\right)\mathbf{J}^\dagger = \mathbf{J}\boldsymbol{e}\otimes(\mathbf{J}\boldsymbol{e})^\dagger = \kappa\boldsymbol{e}\otimes\boldsymbol{e}^\dagger,$$

and $\mathbf{J}\boldsymbol{e}$ and $\boldsymbol{e}$ must be collinear; i.e., $\mathbf{J}\boldsymbol{e} = \lambda\boldsymbol{e}$, where $|\lambda|^2 = \kappa$.

Furthermore, if $\hat{\boldsymbol{\rho}} = \boldsymbol{e}\otimes\boldsymbol{e}^\dagger$ commutes with $\mathbf{J}$, then $\hat{\boldsymbol{\rho}}$ is a congruence eigenmatrix of $\mathbf{J}$.

Proof C.12. If $\mathbf{J}\tilde{\boldsymbol{\rho}} = \tilde{\boldsymbol{\rho}}\mathbf{J}$, then $\tilde{\boldsymbol{\rho}}$ and $\mathbf{J}$ have a common eigenvector (Proof C.9). Therefore,

$$\mathbf{J}\boldsymbol{e} = \lambda\,\boldsymbol{e}$$

where $\boldsymbol{e}$ is the only eigenvector of $\tilde{\boldsymbol{\rho}} = \boldsymbol{e}\otimes\boldsymbol{e}^\dagger$, and (via Proof C.10)

$$\mathbf{J}\tilde{\boldsymbol{\rho}}\mathbf{J}^\dagger = \kappa\tilde{\boldsymbol{\rho}},$$

where $\kappa = |\lambda|^2$.

[7] Although the Mueller matrix associated with $\mathbf{J}$ has four eigenvectors, only two of these describe physical polarization states. For example, two of the eigenvectors of a rotation are complex-valued. Two of the eigenvectors of a Lorentz boost have polarization vectors that lie in the plane that is normal to the boost axis; however, these have $S_0 = 0$, which is not physical.

Finally, if $\mathbf{J}$ is a normal matrix and $\hat{\boldsymbol{\rho}}$ is a congruence eigenmatrix of $\mathbf{J}$, then $\hat{\boldsymbol{\rho}}$ commutes with $\mathbf{J}$.

Proof C.13. If $\mathbf{J}$ is normal, $\hat{\boldsymbol{\rho}} = \boldsymbol{e}\otimes\boldsymbol{e}^\dagger$ and $\mathbf{J}\tilde{\boldsymbol{\rho}}\mathbf{J}^\dagger = \kappa\tilde{\boldsymbol{\rho}}$, then $\mathbf{J}\boldsymbol{e} = \lambda\boldsymbol{e}$ (Proof C.11) and

$$\mathbf{J}\tilde{\boldsymbol{\rho}} = \mathbf{J}\boldsymbol{e}\otimes\boldsymbol{e}^\dagger = \lambda\,\boldsymbol{e}\otimes\boldsymbol{e}^\dagger = \lambda\tilde{\boldsymbol{\rho}}.$$

Also

$$\begin{aligned}
\tilde{\boldsymbol{\rho}}\mathbf{J} &= \boldsymbol{e}\otimes\boldsymbol{e}^\dagger\mathbf{J}\\
&= \boldsymbol{e}\otimes\left(\mathbf{J}^\dagger\boldsymbol{e}\right)^\dagger\\
&= \boldsymbol{e}\otimes(\lambda^*\boldsymbol{e})^\dagger && \text{Proof C.2}\\
&= \lambda\,\boldsymbol{e}\otimes\boldsymbol{e}^\dagger\\
&= \lambda\tilde{\boldsymbol{\rho}}
\end{aligned}$$

Therefore, $\mathbf{J}\tilde{\boldsymbol{\rho}} = \tilde{\boldsymbol{\rho}}\mathbf{J}$.

In conclusion, if $\mathbf{J}$ is normal, then $\hat{\boldsymbol{\rho}}$ is a congruence eigenmatrix of $\mathbf{J}$ if and only if $\hat{\boldsymbol{\rho}}$ commutes with $\mathbf{J}$. Both unitary and Hermitian matrices are normal and their congruence eigenmatrices are explored in the following sections.

C.5.1. *Unitary Matrices*

If $\boldsymbol{\rho}$ commutes with $\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)$, then

$$\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\boldsymbol{\rho}\,\mathbf{R}^\dagger_{\underline{\hat{\boldsymbol{n}}}}(\phi) = \boldsymbol{\rho}\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\mathbf{R}^\dagger_{\underline{\hat{\boldsymbol{n}}}}(\phi) = \boldsymbol{\rho}; \tag{C28}$$

that is, $\boldsymbol{\rho}$ is a congruence eigenmatrix of $\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)$ with associated congruence eigenvalue equal to unity. Conversely, if $\boldsymbol{\rho}$ is a congruence eigenmatrix of $\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)$, then

$$\begin{aligned}
\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\boldsymbol{\rho}\,\mathbf{R}^\dagger_{\underline{\hat{\boldsymbol{n}}}}(\phi) &= \kappa\boldsymbol{\rho}\\
\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\boldsymbol{\rho}\,\mathbf{R}^\dagger_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi) &= \kappa\boldsymbol{\rho}\,\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\\
\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\boldsymbol{\rho} &= \kappa\boldsymbol{\rho}\,\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)
\end{aligned} \tag{C29}$$

and $\boldsymbol{\rho}$ commutes with $\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)$ if $\kappa = 1$. The congruence eigenmatrices of a unitary matrix share a single degenerate congruence eigenvalue, $\kappa = 1$; therefore, any linear combination of the congruence eigenmatrices is also a congruence eigenmatrix of the transformation. Considering the equivalent three-dimensional rotation of the Stokes polarization vector $\underline{\boldsymbol{S}}$; any vector that lies along the axis of rotation remains unchanged by that rotation, including $\underline{\boldsymbol{S}} = 0$ (unpolarized radiation).

C.5.2. *Hermitian Matrices*

Define $b = \cosh\beta$ and $\underline{\boldsymbol{b}} = \underline{\hat{\boldsymbol{m}}}\sinh\beta$ such that

$$\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta) = b\,\boldsymbol{\sigma}_0 + \underline{\boldsymbol{b}}\cdot\underline{\boldsymbol{\sigma}}. \tag{C30}$$

Furthermore, define $\tilde{\boldsymbol{\rho}} = (\tilde{s}_0\,\boldsymbol{\sigma}_0 + \underline{\tilde{\boldsymbol{s}}}\cdot\underline{\boldsymbol{\sigma}})/2$, and note that $|\underline{\tilde{\boldsymbol{s}}}| = \tilde{s}_0$ because $\tilde{\boldsymbol{\rho}}$ is singular / purely polarized. If $\tilde{\boldsymbol{\rho}}$ and $\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)$ commute, then their axes of symmetry are collinear; i.e.,

$$\begin{aligned}\underline{\hat{\boldsymbol{m}}}\times\underline{\tilde{\boldsymbol{s}}} &= 0,\\ \underline{\hat{\boldsymbol{m}}}\cdot\underline{\tilde{\boldsymbol{s}}} &= \pm|\underline{\tilde{\boldsymbol{s}}}| = \pm\tilde{s}_0,\\ \underline{\hat{\boldsymbol{m}}} &= \pm\,\underline{\tilde{\boldsymbol{s}}}/\tilde{s}_0.\end{aligned} \tag{C31}$$

Using Equation (C7),

$$\begin{aligned}\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)\tilde{\boldsymbol{\rho}} &= (b\tilde{s}_0 + \underline{\boldsymbol{b}}\cdot\underline{\tilde{\boldsymbol{s}}})\boldsymbol{\sigma}_0/2 + (b\underline{\tilde{\boldsymbol{s}}} + \tilde{s}_0\underline{\boldsymbol{b}})\cdot\underline{\boldsymbol{\sigma}}/2\\ &= (b\tilde{s}_0 \pm |\underline{\boldsymbol{b}}|\tilde{s}_0)\boldsymbol{\sigma}_0/2 + (b\underline{\tilde{\boldsymbol{s}}} \pm |\underline{\boldsymbol{b}}|\underline{\tilde{\boldsymbol{s}}})\cdot\underline{\boldsymbol{\sigma}}/2\\ &= (b \pm |\underline{\boldsymbol{b}}|)\tilde{s}_0\boldsymbol{\sigma}_0/2 + (b \pm |\underline{\boldsymbol{b}}|)\underline{\tilde{\boldsymbol{s}}}\cdot\underline{\boldsymbol{\sigma}}/2\\ &= (b \pm |\underline{\boldsymbol{b}}|)\tilde{\boldsymbol{\rho}}\\ &= (\cosh\beta \pm \sinh\beta)\tilde{\boldsymbol{\rho}}\\ &= e^{\pm\beta}\tilde{\boldsymbol{\rho}}.\end{aligned} \tag{C32}$$

Therefore,

$$\begin{aligned}\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)\,\tilde{\boldsymbol{\rho}}\,\mathbf{B}^{\dagger}_{\underline{\hat{\boldsymbol{m}}}}(\beta) &= \mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)\,\tilde{\boldsymbol{\rho}}\,\mathbf{B}_{\underline{\hat{\boldsymbol{m}}}}(\beta)\\ &= \mathbf{B}^2_{\underline{\hat{\boldsymbol{m}}}}(\beta)\tilde{\boldsymbol{\rho}}\\ &= e^{\pm 2\beta}\tilde{\boldsymbol{\rho}}.\end{aligned} \tag{C33}$$

Hermitian matrices have two distinct congruence eigenvalues; $\kappa = e^{2\beta}$ when the polarization vector $\underline{\tilde{\boldsymbol{s}}}$ is parallel to the boost axis $\underline{\hat{\boldsymbol{m}}}$, and $\kappa = e^{-2\beta}$ when $\underline{\tilde{\boldsymbol{s}}}$ is anti-parallel to $\underline{\hat{\boldsymbol{m}}}$. Therefore, only these two purely polarized states are congruence eigenmatrices of the transformation.

C.6. *Pure Mueller Matrices*

Various authors (e.g., Barakat 1981; Simon 1982; Cloude 1986) have considered the necessary and sufficient conditions that must be satisfied by the elements of a Mueller matrix for it to have an equivalent Jones matrix representation. Simon (1982) and Cloude (1986) present the most intuitive geometric constraints on a pure Mueller matrix based on its equivalent target coherency matrix (Cloude 1986).

For a Jones matrix $\mathbf{J}$, the equivalent 4×4 Hermitian target coherency matrix $\underline{\underline{\mathbf{N}}}$ is defined via the rank 4 tensor $\boldsymbol{\mathcal{N}} = \mathbf{J}\otimes\mathbf{J}^{\dagger}$, such that Equation (49) yields

$$\begin{aligned}N^{\nu}_{\mu} &= \frac{1}{2}\boldsymbol{\sigma}_{\mu} : \left(\mathbf{J}\otimes\mathbf{J}^{\dagger}\right) : \boldsymbol{\sigma}_{\nu}\\ &= \frac{1}{2}\boldsymbol{\sigma}_{\mu} : \mathbf{J}\left(\boldsymbol{\sigma}_{\nu} : \mathbf{J}^{\dagger}\right)\\ &= \frac{1}{2}k_{\mu}k^{*}_{\nu}.\end{aligned} \tag{C34}$$

Here, $k_{\mu} = \boldsymbol{\sigma}_{\mu} : \mathbf{J}$ are the components of a complex-valued target four-vector $\underline{\boldsymbol{k}}$ (Cloude 1986). Equation (C34) is equivalent to $\underline{\underline{\mathbf{N}}} = \underline{\boldsymbol{k}}\otimes\underline{\boldsymbol{k}}^{\dagger}/2$ and, as noted by Simon (1982), $\underline{\underline{\mathbf{N}}}$ is a scalar multiple of a projection matrix, such that

$$\underline{\underline{\mathbf{N}}}^2 = \mathrm{Tr}\left[\underline{\underline{\mathbf{N}}}\right]\underline{\underline{\mathbf{N}}}. \tag{C35}$$

Proof C.14. Define $\underline{\hat{\boldsymbol{k}}} = \underline{\boldsymbol{k}}/|\underline{\boldsymbol{k}}|$ and the projection $\underline{\underline{\mathbf{P}}} = \underline{\hat{\boldsymbol{k}}}\otimes\underline{\hat{\boldsymbol{k}}}^{\dagger}$ (Eqn. C14), such that

$$\underline{\underline{\mathbf{N}}} = \frac{1}{2}\underline{\boldsymbol{k}}\otimes\underline{\boldsymbol{k}}^{\dagger} = \frac{1}{2}|\underline{\boldsymbol{k}}|^2\underline{\hat{\boldsymbol{k}}}\otimes\underline{\hat{\boldsymbol{k}}}^{\dagger} = \frac{1}{2}|\underline{\boldsymbol{k}}|^2\underline{\underline{\mathbf{P}}}.$$

Note that

$$|\underline{\boldsymbol{k}}|^2 = k_{\mu}k^{*}_{\mu} = 2\mathrm{Tr}\left[\underline{\underline{\mathbf{N}}}\right].$$

Therefore, $\underline{\underline{\mathbf{N}}} = \mathrm{Tr}\left[\underline{\underline{\mathbf{N}}}\right]\underline{\underline{\mathbf{P}}}$ and (via Eqn. C15)

$$\underline{\underline{\mathbf{N}}}^2 = \mathrm{Tr}\left[\underline{\underline{\mathbf{N}}}\right]^2\underline{\underline{\mathbf{P}}}^2 = \mathrm{Tr}\left[\underline{\underline{\mathbf{N}}}\right]^2\underline{\underline{\mathbf{P}}} = \mathrm{Tr}\left[\underline{\underline{\mathbf{N}}}\right]\underline{\underline{\mathbf{N}}}.$$

The mapping between any Mueller matrix $\underline{\underline{\mathbf{M}}}$ and its associated target coherency matrix $\underline{\underline{\mathbf{N}}}$ exploits the symmetry of the transpose used to define the $\tilde{\otimes}$ operator (Eqn. 45). Let $\mathcal{T}$ represent the transpose of covariant tensor indeces, such that $\mathbf{A}\,\tilde{\otimes}\,\mathbf{B} = \mathcal{T}(\mathbf{A}\otimes\mathbf{B})$ and, by symmetry, $\mathbf{A}\otimes\mathbf{B} = \mathcal{T}(\mathbf{A}\,\tilde{\otimes}\,\mathbf{B})$. Using this operator, the rank 4 tensors associated with $\underline{\underline{\mathbf{M}}}$ and $\underline{\underline{\mathbf{N}}}$ are related by $\boldsymbol{\mathcal{U}} = \mathcal{T}(\boldsymbol{\mathcal{N}})$, where $\boldsymbol{\mathcal{U}}$ is defined by Equation (50) and

$$\begin{aligned}\boldsymbol{\mathcal{N}} &= \mathcal{T}(\boldsymbol{\mathcal{U}})\\ &= \mathcal{T}\left(\frac{1}{2}M^{\nu}_{\mu}\boldsymbol{\sigma}_{\mu}\otimes\boldsymbol{\sigma}_{\nu}\right)\\ &= \frac{1}{2}M^{\nu}_{\mu}\boldsymbol{\sigma}_{\mu}\,\tilde{\otimes}\,\boldsymbol{\sigma}_{\nu},\end{aligned} \tag{C36}$$

such that

$$\begin{aligned}N^{\lambda}_{\kappa} &= \frac{1}{2}\boldsymbol{\sigma}_{\kappa} : \boldsymbol{\mathcal{N}} : \boldsymbol{\sigma}_{\lambda}\\ &= \frac{1}{4}M^{\nu}_{\mu}\,\boldsymbol{\sigma}_{\kappa} : \left(\boldsymbol{\sigma}_{\mu}\,\tilde{\otimes}\,\boldsymbol{\sigma}_{\nu}\right) : \boldsymbol{\sigma}_{\lambda}\\ &= \frac{1}{4}M^{\nu}_{\mu}\,\boldsymbol{\sigma}_{\kappa} : \boldsymbol{\sigma}_{\mu}\boldsymbol{\sigma}_{\lambda}\boldsymbol{\sigma}_{\nu}\\ &= \frac{1}{4}M^{\nu}_{\mu}\,\mathrm{Tr}\left[\boldsymbol{\sigma}_{\kappa}\boldsymbol{\sigma}_{\mu}\boldsymbol{\sigma}_{\lambda}\boldsymbol{\sigma}_{\nu}\right].\end{aligned} \tag{C37}$$

Equation (C37) can be written as $\underline{\underline{\mathbf{N}}} = \underline{\underline{\boldsymbol{\mathcal{T}}}} : \underline{\underline{\mathbf{M}}}$, where $\underline{\underline{\boldsymbol{\mathcal{T}}}}$ is the four-dimensional rank 4 tensor defined by

$$T^{\lambda\mu}_{\kappa\nu} = \frac{1}{4}\mathrm{Tr}\left[\boldsymbol{\sigma}_{\kappa}\boldsymbol{\sigma}_{\mu}\boldsymbol{\sigma}_{\lambda}\boldsymbol{\sigma}_{\nu}\right]. \tag{C38}$$

By symmetry, $\underline{\underline{\mathbf{M}}} = \underline{\underline{\boldsymbol{\mathcal{T}}}} : \underline{\underline{\mathbf{N}}}$; that is, $\underline{\underline{\boldsymbol{\mathcal{T}}}}$ is an involution. As for the target coherency matrix derived from a Jones matrix, $\underline{\underline{\mathbf{N}}}$ is Hermitian.

Proof C.15.

$$
\begin{aligned}
&(N_\kappa^\lambda)^* = (T_{\kappa\nu}^{\lambda\mu})^*(M_\mu^\nu)^* \\
&= \frac{1}{4} M_\mu^\nu \,\mathrm{Tr}\left[\boldsymbol{\sigma}_\kappa \boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\lambda \boldsymbol{\sigma}_\nu\right]^* && M_\mu^\nu \in \mathbb{R} \\
&= \frac{1}{4} M_\mu^\nu \,\mathrm{Tr}\left[\left(\boldsymbol{\sigma}_\kappa \boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\lambda \boldsymbol{\sigma}_\nu\right)^\dagger\right] && \mathrm{Tr}\left[\mathbf{A}\right]^* = \mathrm{Tr}\left[\mathbf{A}^\dagger\right] \\
&= \frac{1}{4} M_\mu^\nu \,\mathrm{Tr}\left[\boldsymbol{\sigma}_\nu \boldsymbol{\sigma}_\lambda \boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\kappa\right] && \mathbf{A}\mathbf{B}^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger \\
&= \frac{1}{4} M_\mu^\nu \,\mathrm{Tr}\left[\boldsymbol{\sigma}_\lambda \boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\kappa \boldsymbol{\sigma}_\nu\right] && \mathrm{Tr}\left[\mathbf{A}\mathbf{B}\right] = \mathrm{Tr}\left[\mathbf{B}\mathbf{A}\right] \\
&= N_\lambda^\kappa && \text{Eqn. (C37)}
\end{aligned}
$$

Just as the coherency matrix formed by the outer product of a single Jones vector $\tilde{\boldsymbol{\rho}} = \hat{\boldsymbol{e}} \otimes \hat{\boldsymbol{e}}^\dagger$ corresponds to a purely polarized state (see Eqn. C17), the target coherency matrix formed by the outer product of a single target vector $\underline{\underline{\mathbf{N}}} = \underline{\underline{\boldsymbol{k}}} \otimes \underline{\underline{\boldsymbol{k}}}^\dagger/2$ corresponds to a pure Mueller matrix. Therefore, a Mueller matrix is pure if and only if its target coherency matrix, $\underline{\underline{\mathbf{N}}} = \underline{\underline{\boldsymbol{\mathcal{T}}}} : \underline{\underline{\mathbf{M}}}$ satisfies Equation (C35). Cloude (1986) expressed this definition with the equivalent constraint that a pure Mueller matrix is associated with a target coherency matrix that has only one non-zero eigenvalue, λ. Combined with its associated eigenvector, $\underline{\hat{\boldsymbol{k}}}$, the target vector $\underline{\boldsymbol{k}} = \lambda^{1/2}\underline{\hat{\boldsymbol{k}}}$ can be used to compute (up to arbitrary phase) the Jones matrix associated with a pure Mueller matrix,

$$\mathbf{J} = \frac{1}{2} k_\mu \boldsymbol{\sigma}_\mu. \tag{C39}$$

C.7. *Impure Mueller Matrices*

Some polarimetric transformations are described by impure Mueller matrices that cannot be represented by a linear transformation of the electric field, as described by a Jones matrix. For example, temporal or spectral depolarization occurs when the response of a system varies as a function of time or frequency on characteristic scales that are smaller than the duration or bandwidth over which the Stokes parameters are integrated. In this case, the measured Stokes parameters,

$$\underline{\underline{\boldsymbol{S}}}' = \langle \underline{\underline{\mathbf{M}}}(t,\nu) \rangle \underline{\underline{\boldsymbol{S}}}, \tag{C40}$$

where the angle brackets represent integration over time and/or frequency.

For example, consider stochastic Faraday rotation that causes fluctuations in the position angle $\Delta\Psi$ that are normally distributed with zero mean and standard deviation σ_Ψ. The average Mueller matrix that describes the integrated Stokes parameters at a single radio frequency,

$$
\begin{aligned}
\langle \underline{\underline{\mathbf{R}}}_{\hat{\boldsymbol{v}}}(2\Delta\Psi) \rangle &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \langle\cos 2\Delta\Psi\rangle & -\langle\sin 2\Delta\Psi\rangle & 0 \\ 0 & \langle\sin 2\Delta\Psi\rangle & \langle\cos 2\Delta\Psi\rangle & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned} \tag{C41}
$$

where the angle brackets represent integration over time and $d = \exp(-2\sigma_\Psi^2) < 1$ describes the depolarization of Stokes Q and U. Note that this Mueller matrix is able to depolarize a purely polarized state, something that cannot be done by a linear transformation of the electric field (Proof 4.3).

D. EXAMPLE TRANSFORMATIONS

This section demonstrates some basic transformations, starting with a source described by Stokes parameters

$$\underline{\underline{\boldsymbol{S}}} = [S_0, S_1, 0, 0]^T \tag{D1}$$

and coherency matrix

$$\boldsymbol{\rho} = (S_0\,\boldsymbol{\sigma}_0 + S_1\,\boldsymbol{\sigma}_1)/2. \tag{D2}$$

D.1. *Example Boost*

Consider a boost along $\hat{\underline{\mathbf{1}}} = (1,0,0)^T$,

$$\mathbf{B} = \mathbf{B}_{\hat{\underline{\mathbf{1}}}}(\beta) = \boldsymbol{\sigma}_0 \cosh\beta + \boldsymbol{\sigma}_1 \sinh\beta,$$

such that

$$
\begin{aligned}
\boldsymbol{\rho}' &= \mathbf{B}\,\boldsymbol{\rho}\,\mathbf{B}^\dagger \\
&= \mathbf{B}\,(S_0\,\boldsymbol{\sigma}_0 + S_1\,\boldsymbol{\sigma}_1)\,\mathbf{B}/2 \\
&= S_0 \mathbf{B}\boldsymbol{\sigma}_0\mathbf{B}/2 + S_1\mathbf{B}\boldsymbol{\sigma}_1\mathbf{B}/2.
\end{aligned} \tag{D3}
$$

Using Equation (61), the first term of this equation includes

$$\mathbf{B}_{\hat{\underline{\mathbf{1}}}}(\beta)\,\boldsymbol{\sigma}_0\,\mathbf{B}_{\hat{\underline{\mathbf{1}}}}(\beta) = \mathbf{B}^2_{\hat{\underline{\mathbf{1}}}}(\beta) = \mathbf{B}_{\hat{\underline{\mathbf{1}}}}(2\beta) \tag{D4}$$

The second term includes

$$\begin{aligned}
&\mathbf{B}_{\underline{\hat{\mathbf{1}}}}(\beta)\,\boldsymbol{\sigma}_1\,\mathbf{B}_{\underline{\hat{\mathbf{1}}}}(\beta)\\
&\quad=(\boldsymbol{\sigma}_0\cosh\beta+\boldsymbol{\sigma}_1\sinh\beta)\,\boldsymbol{\sigma}_1\,(\boldsymbol{\sigma}_0\cosh\beta+\boldsymbol{\sigma}_1\sinh\beta)\\
&\quad=(\boldsymbol{\sigma}_0\cosh\beta+\boldsymbol{\sigma}_1\sinh\beta)\,(\boldsymbol{\sigma}_1\cosh\beta+\boldsymbol{\sigma}_0\sinh\beta)\\
&\quad=\boldsymbol{\sigma}_1\cosh^2\beta+\boldsymbol{\sigma}_0\cosh\beta\sinh\beta\\
&\qquad\qquad+\boldsymbol{\sigma}_0\sinh\beta\cosh\beta+\boldsymbol{\sigma}_1\sinh^2\beta\\
&\quad=\boldsymbol{\sigma}_1\left(\cosh^2\beta+\sinh^2\beta\right)+\boldsymbol{\sigma}_0\left(2\sinh\beta\cosh\beta\right)\\
&\quad=\boldsymbol{\sigma}_1\cosh 2\beta+\boldsymbol{\sigma}_0\sinh 2\beta.
\end{aligned}\tag{D5}$$

Therefore,

$$\begin{aligned}
\boldsymbol{\rho}'&=S_0\mathbf{B}_{\underline{\hat{\mathbf{1}}}}(2\beta)/2+S_1\mathbf{B}\boldsymbol{\sigma}_1\mathbf{B}/2\\
&=S_0\left(\boldsymbol{\sigma}_0\cosh 2\beta+\boldsymbol{\sigma}_1\sinh 2\beta\right)/2\\
&\qquad+S_1\left(\boldsymbol{\sigma}_1\cosh 2\beta+\boldsymbol{\sigma}_0\sinh 2\beta\right)/2\\
&=\boldsymbol{\sigma}_0\left(S_0\cosh 2\beta+S_1\sinh 2\beta\right)/2\\
&\qquad+\boldsymbol{\sigma}_1\left(S_0\sinh 2\beta+S_1\cosh 2\beta\right)/2
\end{aligned}\tag{D6}$$

and

$$\begin{aligned}
S_0'&=\boldsymbol{\sigma}_0:\boldsymbol{\rho}'=S_0\cosh 2\beta+S_1\sinh 2\beta\\
S_1'&=\boldsymbol{\sigma}_1:\boldsymbol{\rho}'=S_0\sinh 2\beta+S_1\cosh 2\beta.
\end{aligned}\tag{D7}$$

That is, the Stokes parameters are boosted along the $\underline{\hat{\mathbf{1}}}$ axis with rapidity -2β and Lorentz factor $\gamma=\cosh 2\beta$.

D.2. *Example Rotation*

Consider a rotation about $\underline{\hat{\mathbf{2}}}=(0,1,0)^T$,

$$\mathbf{R}=\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)=\boldsymbol{\sigma}_0\cos\chi+i\boldsymbol{\sigma}_2\sin\chi,$$

such that

$$\begin{aligned}
\boldsymbol{\rho}'&=\mathbf{R}\,\boldsymbol{\rho}\,\mathbf{R}^\dagger\\
&=\mathbf{R}\left(S_0\,\boldsymbol{\sigma}_0+S_1\,\boldsymbol{\sigma}_1\right)\mathbf{R}^\dagger/2\\
&=S_0\mathbf{R}\boldsymbol{\sigma}_0\mathbf{R}^\dagger/2+S_1\mathbf{R}\boldsymbol{\sigma}_1\mathbf{R}^\dagger/2.
\end{aligned}\tag{D8}$$

For the first term of this equation, note that rotations have no effect on $\boldsymbol{\sigma}_0$; e.g.

$$\begin{aligned}
\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\boldsymbol{\sigma}_0\,\mathbf{R}^\dagger_{\underline{\hat{\boldsymbol{n}}}}(\phi)&=\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\mathbf{R}^\dagger_{\underline{\hat{\boldsymbol{n}}}}(\phi)\\
&=\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(\phi)\,\mathbf{R}^{-1}_{\underline{\hat{\boldsymbol{n}}}}(\phi)=\boldsymbol{\sigma}_0
\end{aligned}\tag{D9}$$

Noting that (see Eqn. 62),

$$\mathbf{R}^{-1}_{\underline{\hat{\boldsymbol{n}}}}(\phi)=\mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(-\phi),\tag{D10}$$

the second term includes

$$\begin{aligned}
&\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)\,\boldsymbol{\sigma}_1\,\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(-\chi)\\
&\quad=(\boldsymbol{\sigma}_0\cos\chi+i\boldsymbol{\sigma}_2\sin\chi)\,\boldsymbol{\sigma}_1\,(\boldsymbol{\sigma}_0\cos\chi-i\boldsymbol{\sigma}_2\sin\chi)\\
&\quad=(\boldsymbol{\sigma}_0\cos\chi+i\boldsymbol{\sigma}_2\sin\chi)\,(\boldsymbol{\sigma}_1\cos\chi+\boldsymbol{\sigma}_3\sin\chi)\\
&\quad=\boldsymbol{\sigma}_1\cos^2\chi+\boldsymbol{\sigma}_3\cos\chi\sin\chi\\
&\qquad\qquad+i(-i\boldsymbol{\sigma}_3)\sin\chi\cos\chi+i(i\boldsymbol{\sigma}_1)\sin^2\chi\\
&\quad=\boldsymbol{\sigma}_1\left(\cos^2\chi-\sin^2\chi\right)+2\boldsymbol{\sigma}_3\cos\chi\sin\chi\\
&\quad=\boldsymbol{\sigma}_1\cos 2\chi+\boldsymbol{\sigma}_3\sin 2\chi
\end{aligned}\tag{D11}$$

Therefore,

$$\boldsymbol{\rho}'=S_0\,\boldsymbol{\sigma}_0/2+S_1\left(\boldsymbol{\sigma}_1\cos 2\chi+\boldsymbol{\sigma}_3\sin 2\chi\right)/2\tag{D12}$$

and

$$\begin{aligned}
S_0'&=\boldsymbol{\sigma}_0:\boldsymbol{\rho}'=S_0\\
S_1'&=\boldsymbol{\sigma}_1:\boldsymbol{\rho}'=S_1\cos 2\chi\\
S_3'&=\boldsymbol{\sigma}_3:\boldsymbol{\rho}'=S_1\sin 2\chi.
\end{aligned}\tag{D13}$$

That is, the total intensity is unchanged and the Stokes polarization vector is rotated about the $\underline{\hat{\mathbf{2}}}$ axis by -2χ.

D.3. *Complex-valued Polarization Ellipse*

This section demonstrates the relationship between the geometry of the polarization ellipse (Eqns. [2] through [4]) and the spherical coordinates of the Stokes parameters (Eqn. 7) by deriving the analytic representation of a monochromatic electromagnetic wave (Eqn. 5). First, consider

$$\boldsymbol{e}(t)=\hat{\boldsymbol{e}}_0 re^{i(2\pi\nu t+\phi')}\tag{D14}$$

with polarization defined by the unit Jones vector $\hat{\boldsymbol{e}}_0$, and associated Stokes polarization vector (see Eqn. 7),

$$\underline{\boldsymbol{S}}=r^2\left(\cos 2\chi\cos 2\psi,\ \cos 2\chi\sin 2\psi,\ \sin 2\chi\right).\tag{D15}$$

As in Section 5.3.2, consider the transformation from the original basis to one in which the second component of the electric field vector is zero; i.e.,

$$\boldsymbol{e}''(t)=\mathbf{R}\boldsymbol{e}(t)=\begin{pmatrix}1\\0\end{pmatrix}re^{i(2\pi\nu t+\phi')}.\tag{D16}$$

Invert this using Equation (80), yielding

$$\boldsymbol{e}(t)=\mathbf{R}^{-1}\boldsymbol{e}''(t)=\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)\boldsymbol{e}''(t),\tag{D17}$$

and compute the components of $\hat{\boldsymbol{e}}_0$, starting with the transformation of $\boldsymbol{e}''(t)$ by $\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)$,

$$\begin{aligned}\boldsymbol{e}'(t) &= \mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)\boldsymbol{e}''(t) = (\boldsymbol{\sigma}_0\cos\chi + i\boldsymbol{\sigma}_2\sin\chi)\,\boldsymbol{e}''(t) \\ &= \left[\begin{pmatrix}1&0\\0&1\end{pmatrix}\cos\chi + i\begin{pmatrix}0&1\\1&0\end{pmatrix}\sin\chi\right]\boldsymbol{e}''(t) \\ &= \begin{pmatrix}\cos\chi & i\sin\chi\\ i\sin\chi & \cos\chi\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} re^{i(2\pi\nu t+\phi')} \\ &= \begin{pmatrix}\cos\chi\\ i\sin\chi\end{pmatrix} re^{i(2\pi\nu t+\phi')}.\end{aligned} \tag{D18}$$

The real part of this analytic signal,

$$\boldsymbol{\epsilon}'(t) = \begin{pmatrix} r\cos\chi\cos(2\pi\nu t+\phi') \\ -r\sin\chi\sin(2\pi\nu t+\phi')\end{pmatrix}. \tag{D19}$$

The polarization state of the wave is independent of the absolute phase ϕ'; therefore, choose $\phi' = \phi - \pi/2$ to arrive at Equation (4). Then transform $\boldsymbol{e}'$ by $\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)$ to arrive at Equation (5),

$$\begin{aligned}\boldsymbol{e}(t) &= \mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)\boldsymbol{e}'(t) \\ &= (\boldsymbol{\sigma}_0\cos\psi - i\boldsymbol{\sigma}_3\sin\psi)\,\boldsymbol{e}'(t) \\ &= \left[\begin{pmatrix}1&0\\0&1\end{pmatrix}\cos\psi - i\begin{pmatrix}0&-i\\i&0\end{pmatrix}\sin\psi\right]\boldsymbol{e}'(t) \\ &= \begin{pmatrix}\cos\psi & -\sin\psi\\ \sin\psi & \cos\psi\end{pmatrix}\begin{pmatrix}\cos\chi\\ i\sin\chi\end{pmatrix} re^{i(2\pi\nu t+\phi')} \\ &= \begin{pmatrix}\cos\psi\cos\chi - i\sin\psi\sin\chi\\ \sin\psi\cos\chi + i\cos\psi\sin\chi\end{pmatrix} re^{i(2\pi\nu t+\phi')} \\ &= \boldsymbol{e}_0 e^{i(2\pi\nu t+\phi')},\end{aligned} \tag{D20}$$

where $\boldsymbol{e}_0 = r\hat{\boldsymbol{e}}_0$ as in Equation (6).

This result can be verified by demonstrating that the equivalent congruence transformation of the coherency matrix

$$\boldsymbol{\rho} = \mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)\boldsymbol{\rho}''\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(-\chi)\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(\psi), \tag{D21}$$

yields the original polarization vector $\underline{\boldsymbol{S}}$. Starting with

$$\boldsymbol{\rho}'' = (\boldsymbol{\sigma}_0 + \boldsymbol{\sigma}_1)\,r^2/2, \tag{D22}$$

expand the rotations from the inside outward. Note that rotation of a S_1 polarized state about the S_2 axis has already been derived in Equation (D11), where it is shown that

$$\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)\,\boldsymbol{\sigma}_1\,\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(-\chi) = \boldsymbol{\sigma}_1\cos 2\chi + \boldsymbol{\sigma}_3\sin 2\chi. \tag{D23}$$

As detailed in Appendix C.5.1, the $\boldsymbol{\sigma}_3$ basis matrix is a congruence eigenmatrix of any rotation about the S_3 axis; therefore, it is necessary to consider only the transformation of the $\boldsymbol{\sigma}_1$ basis matrix,

$$\begin{aligned}&\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)\,\boldsymbol{\sigma}_1\,\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(\psi) \\ &\quad= (\boldsymbol{\sigma}_0\cos\psi - i\boldsymbol{\sigma}_3\sin\psi)\,\boldsymbol{\sigma}_1\,(\boldsymbol{\sigma}_0\cos\psi + i\boldsymbol{\sigma}_3\sin\psi) \\ &\quad= (\boldsymbol{\sigma}_0\cos\psi - i\boldsymbol{\sigma}_3\sin\psi)\,(\boldsymbol{\sigma}_1\cos\psi + i(-i\boldsymbol{\sigma}_2)\sin\psi) \\ &\quad= \boldsymbol{\sigma}_1\cos^2\psi + \boldsymbol{\sigma}_2\cos\psi\sin\psi \\ &\qquad\qquad -i(i\boldsymbol{\sigma}_2)\sin\psi\cos\psi - i(-i\boldsymbol{\sigma}_1)\sin^2\psi \\ &\quad= \boldsymbol{\sigma}_1\left(\cos^2\psi - \sin^2\psi\right) + 2\boldsymbol{\sigma}_2\cos\psi\sin\psi \\ &\quad= \boldsymbol{\sigma}_1\cos 2\psi + \boldsymbol{\sigma}_2\sin 2\psi.\end{aligned} \tag{D24}$$

Finally, expand Equation (D21)

$$\begin{aligned}\boldsymbol{\rho} &= \mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)\boldsymbol{\rho}''\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(-\chi)\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(\psi) \\ &= \mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(\chi)\,(\boldsymbol{\sigma}_0 + \boldsymbol{\sigma}_1)\,\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(-\chi)\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(\psi)\,r^2/2 \\ &= \mathbf{R}_{\underline{\hat{\mathbf{3}}}}(-\psi)\,(\boldsymbol{\sigma}_0 + \boldsymbol{\sigma}_1\cos 2\chi + \boldsymbol{\sigma}_3\sin 2\chi)\,\mathbf{R}_{\underline{\hat{\mathbf{3}}}}(\psi)\,r^2/2 \\ &= \frac{r^2}{2}\left[\boldsymbol{\sigma}_0 + (\boldsymbol{\sigma}_1\cos 2\psi + \boldsymbol{\sigma}_2\sin 2\psi)\cos 2\chi + \boldsymbol{\sigma}_3\sin 2\chi\right] \\ &= \frac{r^2}{2}(\boldsymbol{\sigma}_0 + \boldsymbol{\sigma}_1\cos 2\chi\cos 2\psi \\ &\qquad\qquad + \boldsymbol{\sigma}_2\cos 2\chi\sin 2\psi + \boldsymbol{\sigma}_3\sin 2\chi),\end{aligned} \tag{D25}$$

which is consistent with the polarization vector defined in Equation (D15).

D.4. *Transformation to a Circular Basis*

Equation (82) is not the only way to represent a pair of orthonormal circularly-polarized receptors; however, it is the only unimodular unitary matrix that effects the desired cyclic permutation of $\underline{\boldsymbol{S}} = (Q,U,V)^T$ into $\underline{\boldsymbol{S}}' = (V,Q,U)^T$ for all values of Q, U and V (assuming the Stokes V sign convention of IEEE 1983). To see this, consider the following three derivations of Equation (82), each of which provides a slightly different perspective.

D.4.1. *Direct Derivation*

As discussed in Section 3.4, for a left-hand circularly polarized (LCP) wave, the phase of e_y leads that of e_x by 90°. If the LCP wave is also 100% polarized, then

$$e_y(t) = \exp(i\pi/2)e_x(t) = ie_x(t), \tag{D26}$$

which can be expressed as in Equation (12),

$$\boldsymbol{e}_L(t) = \begin{pmatrix}1\\ i\end{pmatrix} z(t). \tag{D27}$$

For RCP, the phase of e_y lags that of e_x by 90° and

$$\boldsymbol{e}_R(t) = \begin{pmatrix} 1 \\ -i \end{pmatrix} z'(t). \tag{D28}$$

Unit receptors that have maximal response to LCP and RCP are given by the normalized Hermitian transposes of the Jones vectors in Equations (D27) and (D28); i.e.,

$$\hat{\boldsymbol{r}}_L = \frac{1}{\sqrt{2}}(1,\, -i) \quad \text{and} \quad \hat{\boldsymbol{r}}_R = \frac{1}{\sqrt{2}}(1,\, i)\,. \tag{D29}$$

The unitary matrix comprised of these row vectors,

$$\mathbf{J}_C = \begin{pmatrix} \hat{\boldsymbol{r}}_L \\ \hat{\boldsymbol{r}}_R \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}, \tag{D30}$$

has a determinant of i, and normalizing by $\sqrt{i}$ yields the unimodular unitary matrix shown in Equation (82). This transformation effects the desired cyclic permutation of the Stokes parameters.

D.4.2. *Indirect Derivation*

Either row vector in Equation (D29) can be multiplied by an arbitrary phase and the pair would continue to be orthonormal. For example, it would be equally reasonable to start with a more symmetric pair of equations,

$$\hat{\boldsymbol{r}}'_L = \hat{\boldsymbol{r}}_L = \frac{1}{\sqrt{2}}(1,\, -i) \tag{D31}$$

and

$$\hat{\boldsymbol{r}}'_R = -i\hat{\boldsymbol{r}}_R = \frac{1}{\sqrt{2}}(-i,\, 1)\,. \tag{D32}$$

The unitary matrix comprised of these orthonormal row vectors,

$$\begin{aligned} \mathbf{R}_C &= \begin{pmatrix} \hat{\boldsymbol{r}}'_L \\ \hat{\boldsymbol{r}}'_R \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}(\boldsymbol{\sigma}_0 - i\boldsymbol{\sigma}_2) = \mathbf{R}_{\underline{\hat{\mathbf{2}}}}(-\pi/4), \end{aligned} \tag{D33}$$

rotates the polarization vector $\underline{\boldsymbol{S}} = (Q, U, V)^T$ around the S_2 axis by 90°, yielding $\underline{\boldsymbol{S}}' = (V, U, -Q)^T$.

To arrive at the desired cyclic permutation, note that a 90° rotation about the S_1 axis transforms $(V, U, -Q)^T$ into $(V, Q, U)^T$. Therefore,

$$\begin{aligned} \mathbf{R}_b &= \mathbf{R}_{\underline{\hat{\mathbf{1}}}}(-\pi/4)\mathbf{R}_{\underline{\hat{\mathbf{2}}}}(-\pi/4) \\ &= \begin{pmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \\ &= \frac{\sqrt{-i}}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} = \frac{1}{\sqrt{2i}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}. \end{aligned} \tag{D34}$$

D.4.3. *Geometric Derivation*

Rather than starting with the electric field, consider the three-dimensional rotation of the Stokes polarization vector that effects the desired cyclic permutation. The eigenvectors of this rotation lie on the axis defined by $Q = U = V$; therefore, choose $\underline{\hat{\boldsymbol{n}}} = (1, 1, 1)^T/\sqrt{3}$ and note that a 120° rotation about this axis rotates Stokes V into S'_1, Stokes Q into S'_2, and Stokes U into S'_3. The equivalent unitary transformation of the electric field,

$$\begin{aligned} \mathbf{R}_b = \mathbf{R}_{\underline{\hat{\boldsymbol{n}}}}(-\pi/3) &= \frac{1}{2}\left(\boldsymbol{\sigma}_0 - i\sum_{k=1}^{3}\boldsymbol{\sigma}_k\right) \\ &= \frac{1}{2}\begin{pmatrix} 1-i & -(1+i) \\ 1-i & 1+i \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\pi/4} & -e^{i\pi/4} \\ e^{-i\pi/4} & e^{i\pi/4} \end{pmatrix} \\ &= \frac{1}{\sqrt{2i}}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}. \end{aligned} \tag{D35}$$

E. STOKES PARAMETERS AS INTENSITY DIFFERENCES

Currently, at frequencies above the microwave region of the electromagnetic spectrum, it is not possible to directly sample the electric field and therefore not possible to simultaneously compute all four Stokes parameters as in Equations (17) through (20). Rather, only S_0 and S_1 can be directly measured because they are the sums and differences of the flux densities (or intensities) of the radiation after passing through oppositely polarized filters.

In the basis defined by the Cartesian coordinates introduced in Section 2, the Stokes polarization vector $\underline{\boldsymbol{S}}_q = (Q, U, V)^T$, and Equation (18) yields

$$S_1 = |e_x|^2 - |e_y|^2 = I_x - I_y = Q, \tag{E1}$$

where I_x and I_y are the intensities of the electromagnetic radiation after passing through a linearly polarized filter with its transmission axis oriented along the x and y axes, respectively.

In a basis that is formed by rotating the original basis by 45° around the line of sight (e.g., the x' and y' axes of Section 2 when $\psi = \pi/4$), the Stokes polarization vector $\underline{\boldsymbol{S}}_u = (U, -Q, V)^T$, and Equation (18) yields

$$S'_1 = |e'_x|^2 - |e'_y|^2 = I'_x - I'_y = U, \tag{E2}$$

where I'_x and I'_y are the intensities measured after passing through a linearly polarized filter with its transmission axis oriented along the x' and y' axes, respectively.

In a basis defined by receptors with orthogonal senses of circular polarization, as discussed in Section 5.4.1, the Stokes polarization vector $\underline{\boldsymbol{S}}_v = (V, Q, U)^T$, and Equation (18) yields

$$S_1^c = |e_l|^2 - |e_r|^2 = I_l - I_r = V, \tag{E3}$$

where I_l and I_r are the intensities after passing through filters that pass left and right circularly polarized radiation, respectively.

In all three cases, the total intensity S_0 is given by the sum of the intensities of the orthogonal polarizations, which is independent of the basis in which it is measured. By measuring the six intensities that appear in Equations (E1) through (E3), which correspond to the six special cases plotted in Figure 3, all four Stokes parameters can be measured.

F. REFINEMENTS TO EXISTING DEFINITIONS

This article introduces some minor refinements and corrections to existing definitions that enhance their clarity and accuracy. First, in Stokes (1852), the ellipticity angle χ varies between $-90°$ and $90°$, as inferred from "*the numerical value of [χ] being supposed not to lie beyond the limits 0 and 90°*" and "*polarization is right-handed or left-handed according to the sign of [χ].*" (The "numerical value" is understood to mean the "absolute value.") However, as noted in Section 7.3, allowing $|\chi| > 45°$ leads to model degeneracy; therefore $|\chi| \leq 45°$ in this work.

The IEEE (1983) standard defines the axial ratio r' as the "*ratio of the major to minor axes of a polarization ellipse*" that "*carries a sign that is taken as plus if the sense of polarization is right-handed and minus if it is left-handed.*" The caption to Figure 1 of this article defines a ratio between semi-minor and semi-major axis $r \equiv \tan\chi = -1/r'$ that is positive for left-handed polarization. The sign of r' is negated and, relative to r, it is effectively inverted in the IEEE (1983) definition of the Poincaré sphere, where "*the latitude is twice the angle whose cotangent is the negative of the axial ratio of the polarization ellipse.*" That is, the latitude, $2\chi' = -2\cot^{-1} r' = 2\tan^{-1} r = 2\chi$; therefore, this article and the IEEE standard arrive at mutually consistent definitions of latitude in the Poincaré sphere. However, though only briefly mentioned in this article, the ratio r is preferred because it does not approach infinity for any polarized states and it requires no negation when it is related to latitude in the Poincaré sphere.

When defining the axis-angle representations of unitary and Hermitian matrices, Britton (2000) wrote that unitary transformations rotate the Stokes polarization vector about $\hat{\underline{\boldsymbol{n}}}$ by 2ϕ, and Hermitian transformations boost the Stokes four-vector along $\hat{\underline{\boldsymbol{m}}}$ by 2β. These definitions describe the inverses of the transformations defined in this article and verified in Appendix D.

G. DOWN-CONVERSION

By the Nyquist Theorem, a signal must be discretely sampled at a rate equal to twice its bandwidth in order to completely represent its information content. Therefore, subject to the finite recording rate of digital observatory equipment, a radio astronomical signal must be constrained to a limited portion of the radio spectrum. The intermediate process by which the signal from the receiver is band-limited and made ready for discrete sampling is known as down-conversion.

Consider the incoming radio signal, $x(t)$, and its Fourier transform, $X(\nu)$. The band-limited signal of interest, $x_b(t)$, is parameterized by its centre frequency, ν_0, and bandwidth, $\Delta\nu$. Baseband down-conversion is the process by which the spectral information originally contained in the range $[\nu_0 - \Delta\nu/2,\, \nu_0 + \Delta\nu/2]$ is shifted down to $[0,\, \Delta\nu]$.

The spectral information is shifted to baseband by demodulating or mixing the radio frequencies (RF) with a local oscillator (LO). This is equivalent to multiplying the signal, $x(t)$, with a pure tone, $l(t) = a\cos(2\pi\nu_l t + \phi)$. By application of the convolution theorem, and with reference to Table 2, mixing may also be understood as a convolution with a pair of complex-valued delta functions in the frequency domain. This understanding proves useful in the following sections.

In addition to mixing to lower frequencies, the signal must also be band-limited before analog-to-digital conversion. Otherwise, power from frequencies higher than the Nyquist frequency will be reflected back into the band of interest, a pollution known as aliasing. The ideal low-pass filter is represented by the rectangle function (see Table 2) so that a bandpass filter with centre frequency, ν_0, and bandwidth, $\Delta\nu$, is given by

$$\Pi\left(\frac{|\nu| - \nu_0}{\Delta\nu}\right). \tag{G1}$$

Note that the absolute value of ν in the first term of this equation creates a bandpass window at both positive and negative frequency values.

Down-conversion therefore refers to the combined operation of mixing and band-limiting. The following two subsections describe in detail two commonly used methods of down-conversion: single-sideband (SSB) and dual-sideband (DSB). These are also represented graphically in Figures 6 through 8. The process of down-conversion is performed separately and (ideally) identically on each of the two orthogonal senses of polarization from the receiver feed.

G.1. *Single-sideband Down-conversion*

During single-sideband down-conversion (SSB), the signal of interest, $x(t)$, is first bandpass filtered, producing $x_b(t)$ where $X_b(\nu) = X(\nu)\Pi((|\nu| - \nu_0)/\Delta\nu)$. The band-limited signal is then mixed with a LO with frequency, ν_1, producing $x_m(t) = x_b(t)\cos(2\pi\nu_1 t + \theta)$, where ν_1 is set to either $\nu_0 - \Delta\nu/2$ (upper-sideband, shown in Figure 6), producing $x_u(t)$; or $\nu_0 + \Delta\nu/2$ (lower-sideband, shown in Figure 7), producing $x_l(t)$. After another stage of low-pass filtering, either $x_u(t)$ or $x_l(t)$ is digitally sampled at the Nyquist rate, $2\Delta\nu$. Compared to $x_u(t)$, the two halves of the spectrum in $x_l(t)$ are swapped, which is equivalent to negating the direction on the frequency axis and taking the complex conjugate of the spectrum.

During playback, the analytic signal associated with $x(t)$ may be formed in practice by taking the real-to-complex Fast Fourier Transform (FFT), followed by the complex-to-complex inverse FFT. Most real-to-complex FFT implementations automatically omit the redundant negative frequencies ($X(-\nu) = X^*(\nu)$) from their output, implicitly producing the analytic signal. Since many signal processing operations (such as phase-coherent dispersion removal) are performed in the Fourier domain, the cost of calculating the analytic signal is transparent.

G.2. *Dual-sideband Down-conversion*

During dual-sideband down-conversion (DSB, see Figure 8), also known as quadrature mixing, the voltages from the receiver are split equally into two signal paths. One signal is mixed with a local oscillator, producing

$$i(t) = x(t)\cos(2\pi\nu_0 t). \tag{G2}$$

The other signal is mixed with the same local oscillator phase-shifted by 90°,

$$q(t) = x(t)\sin(2\pi\nu_0 t). \tag{G3}$$

Both $i(t)$ and $q(t)$ are low-pass filtered with a cutoff frequency of $\nu_c = \Delta\nu/2$, producing $i_b(t) = i(t) * \mathrm{sinc}(\pi\Delta\nu t)$ and $q_b(t) = q(t) * \mathrm{sinc}(\pi\Delta\nu t)$. The low-pass filtered signals are then digitally sampled at the Nyquist rate of $2\nu_c = \Delta\nu$. The signals, $i_b(t)$ and $q_b(t)$ are known as the in-phase and quadrature components, respectively, of $x(t)$ with respect to ν_0.

During playback, the analytic signal associated with $x(t)$ is given by

$$\begin{aligned} z_b(t) &= i_b(t) + iq_b(t) \\ &= [x(t)\cos(2\pi\nu_0 t) + ix(t)\sin(2\pi\nu_0 t)] * \mathrm{sinc}(\pi\Delta\nu t) \\ &= [x(t)e^{2\pi i\nu_0 t}] * \mathrm{sinc}(\pi\Delta\nu t). \end{aligned} \tag{G4}$$

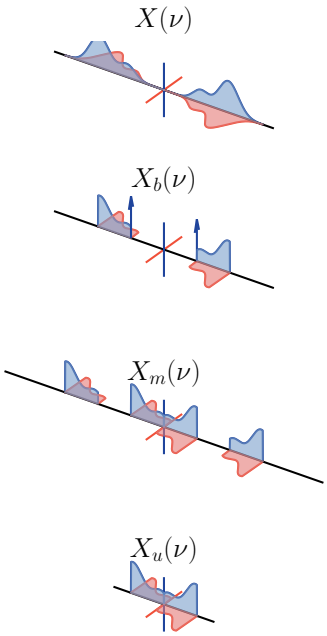

Figure 6. Upper-sideband down-conversion. From the top, the real-valued signal, $x(t)$ has a complex-valued spectrum with conjugate symmetry, such that $X(-\nu) = X^*(\nu)$, as depicted by $\mathrm{Re}[X(\nu)]$ in blue and $\mathrm{Im}[X(\nu)]$ in red. The signal is bandpass filtered and the band-limited signal, $X_b(\nu)$, is mixed with a local oscillator with frequency $\nu_0 - \Delta\nu/2$, as depicted by a pair of blue delta functions. The resulting signal, $X_m(\nu)$, is low-pass filtered, producing $X_u(\nu)$. For each spectrum, the frequency axis is drawn in black, and the real and imaginary parts of each complex value are projected along the blue and red axes (respectively) depicted at the origin (where $\nu = 0$).

In the Fourier domain,

$$Z_b(\nu) = X(\nu + \nu_0)\Pi(\nu/\Delta\nu). \tag{G5}$$

That is, the spectrum of $z_b(t)$ is equivalent to the band-limited portion of $x(t)$ centred at ν_0. The negative frequency components, centred at $-\nu_0$, have been suppressed by low-pass filtering, forming the analytic signal associated with $x(t)$.

Note that changing the sign of the 90° phase shift in Equation (G3) changes the sign of $q(t)$, which results in complex conjugation of $z_b(t)$. This is equivalent to modulating $x(t)$ with a local oscillator that has frequency $-\nu_0$, which shifts the negative half of the spectrum of $X(\nu)$ to baseband. As in the case of lower-sideband down-conversion, the resulting spectrum is complex conjugated and the frequency axis is negated.

During SSB down-conversion, bandpass filtering is performed before mixing, necessitating a filter that is tunable over the range of frequencies of interest. In contrast, the low-pass filter used for DSB down-conversion remains constant regardless of the desired observing fre-

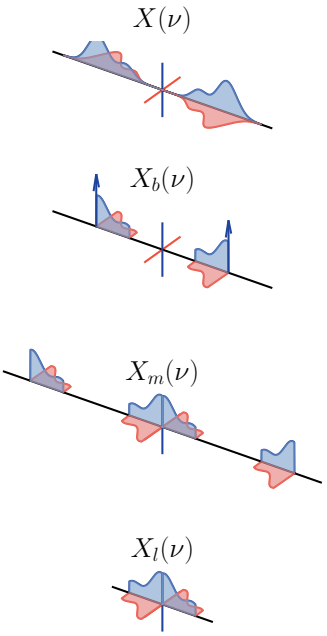

Figure 7. Lower-sideband down-conversion. As for upper-sideband except the band-limited signal, $X_b(\nu)$, is mixed with a local oscillator with frequency $\nu_0 + \Delta\nu/2$. The resulting signal, $X_m(\nu)$, is low-pass filtered, producing $X_l(\nu)$.

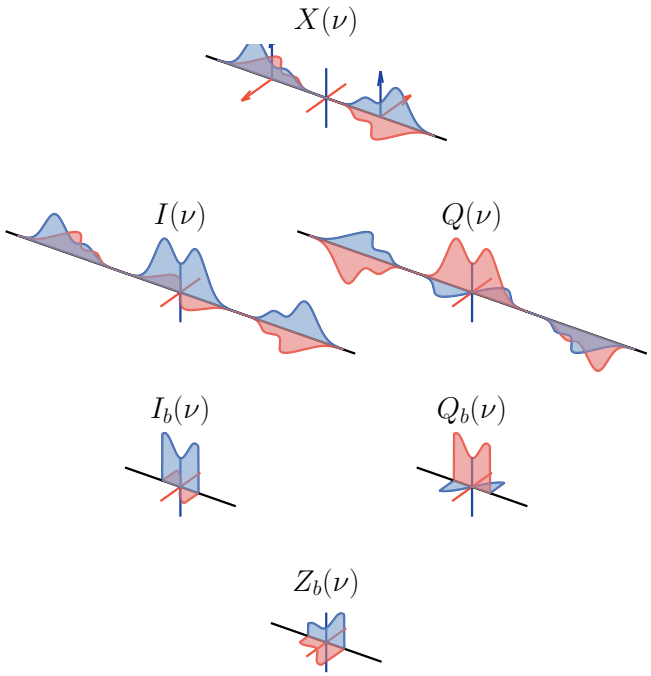

Figure 8. Dual-sideband down-conversion (DSB). The real-valued signal, $X(\nu)$, is split before mixing with a local oscillator with frequency ν_0 (blue delta functions) and a 90° phase-shifted local oscillator (red delta functions), yielding the in-phase and quadrature components, $\mathrm{I}(\nu)$ and $\mathrm{Q}(\nu)$, respectively. Each of the signals are low-pass filtered, yielding $\mathrm{I}_b(\nu)$ and $\mathrm{Q}_b(\nu)$ with bandwidth $\Delta\nu/2$. The complex signal, $\mathrm{Z}_b(\nu)=\mathrm{I}_b(\nu)+i\mathrm{Q}_b(\nu)$, is the analytic signal associated with $X_u(\nu)$, which is depicted in Figure 6.

quency. For this reason, DSB is often the more economical means of down-conversion.

G.3. *Nyquist Zones*

In the examples so far, the down-converted signal was placed in the first Nyquist zone, which spans from $-\Delta\nu$ to $\Delta\nu$ in the case of single-sideband down-conversion, and from $-\Delta\nu/2$ to $\Delta\nu/2$ for dual-sideband down-conversion. The process of discretely sampling the analog signal $x(t)$ at uniformly-spaced instances in time is equivalent to convolving the band-limited spectrum with an infinite series of delta functions separated by the sampling frequency. An example of sampling an analog signal after upper-sideband down-conversion to the first Nyquist zone is depicted in Figure 9.

It is also possible to place the down-converted and band-limited analog signal in a region of spectrum offset from the first Nyquist zone. If the spectrum is offset by $N\Delta\nu$, where N is an integer, then the spectrum is said to be placed in the $(N+1)^{\mathrm{th}}$ Nyquist zone, and the digitized signal will have frequencies that either increase or decrease linearly across the spectrum. An example of sampling an analog signal after upper-sideband down-conversion to the second ($N = 1$) Nyquist zone is depicted in Figure 10. With respect to the digitized signal depicted in Figure 9, the digitized signal in Figure 10 is shifted by $\Delta\nu$, which is equivalent to what would be obtained by digitizing an analog signal af-

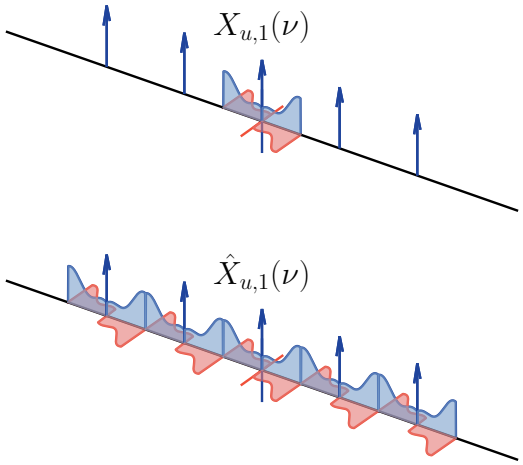

Figure 9. Down-conversion to the first Nyquist zone before sampling. The band-limited analog signal, $X_{u,1}(\nu)$ is depicted following upper-sideband down-conversion, along with five of the infinite series of delta functions separated by the sampling frequency. The digitized signal $\hat{X}_{u,1}(\nu)$ is the convolution of the analog signal and the sampling function.

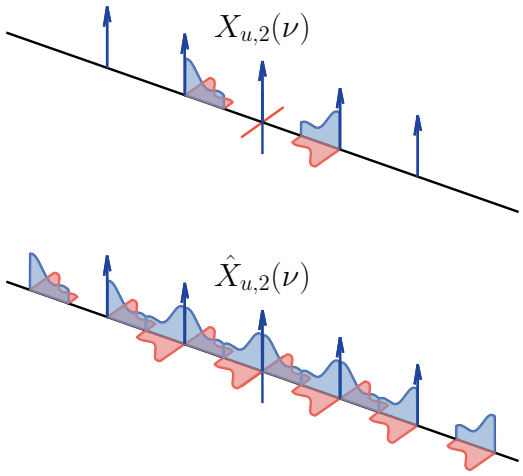

Figure 10. Down-conversion to the second Nyquist zone before sampling. As for Figure 9, except that the analog signal has been down-converted to occupy the second Nyquist zone before analog-to-digital conversion.

ter lower-sideband down-conversion to the first Nyquist zone. That is, the spectrum is complex conjugated and the frequency axis is negated.

H. REFLECTION AS TURNING OVER THE RECEIVER

Reflection can also be modeled as turning over the receiver, which is equivalent to rotating the reference frame depicted in Figure 1 by 180° about a transverse axis in the x-y plane. This negates the z axis and results in a two-dimensional reflection in the x-y plane through the rotation axis. Let $\hat{\underline{\boldsymbol{a}}} = (\cos\alpha, \sin\alpha, 0)^T$ define the three-dimensional rotation axis in the physical space of Figure 1 and use Rodrigues' Rotation Formula to express the 3×3 matrix for a 180° rotation about $\hat{\underline{\boldsymbol{a}}}$.

$$\begin{pmatrix} 2a_x^2 - 1 & 2a_x a_y & 0 \\ 2a_x a_y & 2a_x^2 - 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{H1}$$

In the x-y plane, this transformation reduces to

$$\mathbf{W} = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix} = \cos 2\alpha\, \boldsymbol{\sigma}_1 + \sin 2\alpha\, \boldsymbol{\sigma}_2, \tag{H2}$$

which is a linear combination of the two-dimensional reflection basis matrices defined by the two real-valued Pauli matrices.[8] A reflection in the x-y plane negates the *handedness*[9] of the two-dimensional basis, which can also be reversed by swapping the cables used to propagate the orthogonally polarized signals. In van Straten et al. (2010), the *feed hand* defines the handedness of the receptor basis and the reflection is performed through the axis defined by the *symmetry angle*.

The polar decomposition of a two-dimensional reflection through the axis defined by $(\cos\alpha, \sin\alpha)$ yields a unitary matrix parameterized by $\hat{\underline{\boldsymbol{n}}} = (\cos 2\alpha\ \sin 2\alpha, 0)^T$,

$$\mathbf{W}(\alpha) = -i\, \mathbf{R}_{\hat{\underline{\boldsymbol{n}}}}(\pm\pi/2). \tag{H3}$$

Proof H.1. Note that $|\mathbf{W}|^{1/2} = \pm i$ and

$$\begin{aligned} \mathbf{W}(\alpha) &= -i\, \mathbf{R}_{\hat{\underline{\boldsymbol{n}}}}(\pi/2) \\ &= -i \left(\boldsymbol{\sigma}_0 \cos\frac{\pi}{2} + i\, \underline{\boldsymbol{n}} \cdot \underline{\boldsymbol{\sigma}} \sin\frac{\pi}{2} \right) \\ &= \underline{\boldsymbol{n}} \cdot \underline{\boldsymbol{\sigma}} \\ &= \cos 2\alpha\, \boldsymbol{\sigma}_1 + \sin 2\alpha\, \boldsymbol{\sigma}_2 \end{aligned}$$

Congruence transformation by $\mathbf{W}(\alpha)$ rotates $\underline{\boldsymbol{S}}$ by 180° about the $\hat{\underline{\boldsymbol{n}}}$ axis defined by α, which negates both the ellipticity angle χ and the position angle ψ.

In a basis defined by linearly-polarized receptors, the nominal axis of symmetry has a position angle of 45°. The reflection defined by $\mathbf{W}(\pi/4) = \boldsymbol{\sigma}_2$ swaps e_x and e_y. With reference to Equations (18) through (20), swapping e_x and e_y negates Stokes Q and V. This negation is equivalent to a ±180° rotation of the Stokes polarization vector about $\hat{\underline{\boldsymbol{n}}} = (0, 1, 0)^T$, which is also the result of congruence transformation by $\mathbf{W}(\pi/4)$.

In the circular basis, the nominal axis of symmetry has a position angle of 0°, and $\mathbf{W}(0) = \boldsymbol{\sigma}_1$ negates e_y. Negating either component of the electric field vector reverses the signs of both Stokes U and V, which is equivalent to the ±180° rotation of the Stokes polarization vector about $\hat{\underline{\boldsymbol{n}}} = (1, 0, 0)^T$ under congruence transformation by $\mathbf{W}(0)$.

I. FUNDAMENTAL NUMERICAL INSTABILITY

This section considers two forms of instability that arise when fitting a mathematical model of the instrumental response to experimental data. First, model degeneracy arises when there is no unique solution to the measurement equation; in this case, the Hessian matrix is singular and matrix inversion fails. Second, when

[8] Both $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ satisfy the criteria of a 2×2 reflection matrix, which must be orthogonal ($\mathbf{W}\mathbf{W}^T = \mathbf{W}^T\mathbf{W} = \boldsymbol{\sigma}_0$) and have eigenvalues, $\lambda = \pm 1$, and determinant, $|\mathbf{W}| = -1$.

[9] Strictly, handedness is a property of only three-dimensional vector spaces; this property is formally known as the orientation of the ordered basis.

one or more model parameters are highly collinear, the Hessian matrix is ill-conditioned, matrix inversion is numerically unstable, and the uncertainties of the collinear model parameters are inflated.

The model under consideration relates the observed Stokes parameters $\underline{\underline{\boldsymbol{S}}}'$ to the unknown Stokes parameters $\underline{\underline{\boldsymbol{S}}}$ intrinsic to the source by a Mueller matrix $\underline{\underline{\mathbf{M}}}$ that describes the unknown instrumental response via

$$\underline{\underline{\boldsymbol{S}}}'(\underline{\underline{\mathbf{X}}};\boldsymbol{\alpha}) = \underline{\underline{\mathbf{M}}}\,\underline{\underline{\mathbf{X}}}\,\underline{\underline{\boldsymbol{S}}}. \tag{I1}$$

In this measurement equation, $\underline{\underline{\mathbf{X}}}$ is the known constraining transformation, or matrix argument, and $\boldsymbol{\alpha}$ is the set of model parameters that describe both $\underline{\underline{\mathbf{M}}}$ and $\underline{\underline{\boldsymbol{S}}}$. The matrix argument $\underline{\underline{\mathbf{X}}}$ can be the projection between the receptors and the celestial sphere (Section 5.5), which can vary with time owing to the rotation of the Earth or mechanical rotation of the feed horn. Faraday rotation (Section 5.6) may also serve as the matrix argument under the assumption that the intrinsic Stokes parameters do not vary with radio frequency ν, or are a known function of ν (e.g., Edwards & Stappers 2004).

Both forms of numerical instability arise when solving Equation (I1) through variation of $\boldsymbol{\alpha}$. Model degeneracy occurs when unknown components of $\underline{\underline{\mathbf{M}}}$ commute with $\underline{\underline{\mathbf{X}}}$, and parameter collinearity arises when unknown components of $\underline{\underline{\boldsymbol{S}}}$ are eigenvectors of $\underline{\underline{\mathbf{X}}}$.

I.1. *Commuting with the Matrix Argument*

Appendix B of van Straten (2004) proves that no unique solution to the polarization measurement equation can be derived when only unknown sources are observed at multiple parallactic angles. Here, the proof is repeated and extended to include impure Mueller matrices (Appendix C.7). Consider

$$\underline{\underline{\boldsymbol{S}}}' = \underline{\underline{\mathbf{M}}}\,\underline{\underline{\mathbf{R}}}(\Phi)\underline{\underline{\boldsymbol{S}}} \tag{I2}$$

where $\underline{\underline{\mathbf{R}}}(\Phi)$ represents a rotation about the line of sight by the parallactic angle Φ.

Given $\underline{\underline{\mathbf{M}}}$ and $\underline{\underline{\boldsymbol{S}}}$ that satisfy this equation for all Φ, it is possible to define a family of solutions, $\underline{\underline{\mathbf{M}}}_u = \underline{\underline{\mathbf{M}}}\,\underline{\underline{\mathbf{U}}}^{-1}$ and $\underline{\underline{\boldsymbol{S}}}_u = \underline{\underline{\mathbf{U}}}\,\underline{\underline{\boldsymbol{S}}}$, where $\underline{\underline{\mathbf{U}}}$ is any matrix that commutes freely with $\underline{\underline{\mathbf{R}}}(\Phi)$ for all values of Φ, such that

$$\begin{aligned}\underline{\underline{\boldsymbol{S}}}' &= \underline{\underline{\mathbf{M}}}_u\,\underline{\underline{\mathbf{R}}}(\Phi)\underline{\underline{\boldsymbol{S}}}_u \\ &= \underline{\underline{\mathbf{M}}}\,\underline{\underline{\mathbf{U}}}^{-1}\underline{\underline{\mathbf{R}}}(\Phi)\underline{\underline{\mathbf{U}}}\,\underline{\underline{\boldsymbol{S}}} \\ &= \underline{\underline{\mathbf{M}}}\,\underline{\underline{\mathbf{U}}}^{-1}\underline{\underline{\mathbf{U}}}\,\underline{\underline{\mathbf{R}}}(\Phi)\underline{\underline{\boldsymbol{S}}} \\ &= \underline{\underline{\mathbf{M}}}\,\underline{\underline{\mathbf{R}}}(\Phi)\underline{\underline{\boldsymbol{S}}}.\end{aligned} \tag{I3}$$

Pure Mueller matrices that commute with $\underline{\underline{\mathbf{R}}}(\Phi)$ include rotations about the Stokes V axis, and Lorentz boosts along the Stokes V axis. Impure Mueller matrices that commute with $\underline{\underline{\mathbf{R}}}(\Phi)$ can be described using equation (46) of Lu & Chipman (1996, hereafter LC96),

$$\underline{\underline{\mathbf{M}}}_\Delta \equiv \begin{pmatrix} 1 & \underline{\mathbf{0}}^T \\ \underline{\boldsymbol{P}}_\Delta & \underline{\underline{\mathbf{M}}}_\Delta \end{pmatrix}, \tag{I4}$$

where $\underline{\boldsymbol{P}}_\Delta = (P_1, P_2, P_3)^T$ is the polarizance vector that describes the conversion of total intensity to polarized flux, $|\underline{\boldsymbol{P}}_\Delta| \leq 1$, $\underline{\mathbf{0}} = (0,0,0)^T$, and $\underline{\underline{\mathbf{M}}}_\Delta$ is the 3×3 symmetric depolarizer matrix. This matrix can be diagonalized by a similarity transformation and written in the form of equation (43) of LC96,

$$\underline{\underline{\mathbf{M}}}_\Delta = \underline{\underline{\mathbf{R}}}^{-1}\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}\underline{\underline{\mathbf{R}}}, \quad |a|,|b|,|c| \leq 1. \tag{I5}$$

where $\underline{\underline{\mathbf{R}}}$ is a 3×3 rotation matrix. In this form, it can be seen that $\underline{\underline{\mathbf{M}}}_\Delta$ commutes with $\underline{\underline{\mathbf{R}}}(\Phi)$ if

- only Stokes V is polarized by $\underline{\boldsymbol{P}}_\Delta$; and/or
- $\underline{\underline{\mathbf{R}}}$ is a rotation about the Stokes V axis, and
  - Stokes Q and U are equally depolarized by $\underline{\underline{\mathbf{M}}}_\Delta$, and/or
  - Stokes V is depolarized by $\underline{\underline{\mathbf{M}}}_\Delta$.

For linearly-polarized receptors, only Stokes V is polarized if $\underline{\boldsymbol{P}}_\Delta = (0,0,P_3)^T$. Furthermore, Stokes Q and U are equally depolarized by $\underline{\underline{\mathbf{M}}}_\Delta$ when $a = b$ (as for the case of stochastic Faraday rotation described in Appendix C.7).

Note that depolarization of Stokes V includes negation ($c = -1$) and, as described in Section 5.1, no linear transformation of the electric field can negate the sign of only Stokes V. For example, if the instrumental response is modeled using Jones matrices, then negation of Stokes V in the unknown $\underline{\underline{\boldsymbol{S}}}$ can be compensated only by rotating the reference frame. However, a $\pm 180°$ rotation that negates Stokes V must also negate the position angle and derived quantities such as the Faraday rotation measure. Therefore, the Stokes V sign ambiguity can be eliminated by observing a source for which the sign of the position angle and/or rotation measure is known.

Owing to commutation, there is no unique solution to Equation (I2) and other constraints or assumptions must be introduced to constrain the degenerate dof. Similar degeneracy will arise whenever the experimental constraints include only observations of unknown sources of radiation as a function of a matrix argument with a fixed axis of symmetry.

When modeling variations of the observed Stokes parameters as a function of a matrix argument with a variable axis of symmetry, it is no longer possible for an unknown component of $\underline{\underline{\mathbf{M}}}$ to commute with all values of $\underline{\underline{\mathbf{X}}}$, and the degeneracy is eliminated. However, as described in the following section, even when the axis of symmetry of $\underline{\underline{\mathbf{X}}}$ is variable, unknown model parameters can be highly collinear and cause numerical instability.

I.2. *Eigenvectors of the Matrix Argument*

In some experiments, the observed Stokes parameters are related to the intrinsic Stokes parameters by a measurement equation that includes a matrix argument with a variable axis of symmetry. Such a measurement equation may have no fundamental degeneracy; however, collinearity between model parameters can arise when one or more unknown components of the polarization states used as constraints are eigenvectors of the matrix argument.

For example, when observations are made over a wide range of hour angles with a fixed dipole array, the geometric projection transformation,

$$\underline{\underline{\mathbf{P}}} \simeq \underline{\underline{\mathbf{B}}}_L(l,m)\underline{\underline{\mathbf{R}}}(\Phi), \tag{I6}$$

where $\underline{\underline{\mathbf{B}}}_L(l,m)$ approximates the foreshortening of the projected receptors, a direction-dependent Lorentz boost with variable axis of symmetry in the Q–U plane, and $\underline{\underline{\mathbf{R}}}(\Phi)$ models the rotation of the observatory about the line of sight by the parallactic angle Φ, a rotation with fixed symmetry along the Stokes V axis. The projection has a symmetry axis that varies with time and therefore there is no fundamental degeneracy.

However, the Stokes vector, $\underline{\underline{\boldsymbol{V}}} = [0,0,0,V]^T$ has a polarization vector that is parallel to the symmetry axis of $\underline{\underline{\mathbf{R}}}$ and perpendicular to the symmetry axis of $\underline{\underline{\mathbf{B}}}_L$. Therefore, it is an eigenvector of $\underline{\underline{\mathbf{P}}}$ with associated eigenvalue $\lambda = 1$; i.e.,

$$\underline{\underline{\mathbf{P}}}\,\underline{\underline{\boldsymbol{V}}} = \underline{\underline{\boldsymbol{V}}}. \tag{I7}$$

Consequently, when the model parameters include unknown circular polarization intrinsic to the sources used as constraints, the Stokes V components are highly covariant with the unknown instrumental boost along the Stokes V axis, $\underline{\underline{\mathbf{B}}}_V(\beta)$, causing the Hessian matrix to be ill-conditioned.

To demonstrate this, first consider the typical Gauss–Newton approximation of the Hessian matrix $\mathbf{H}$. Here, second-order derivative terms are ignored and

$$\mathbf{H} \simeq 2\boldsymbol{\Upsilon}^T\boldsymbol{\Upsilon}, \tag{I8}$$

where $\boldsymbol{\Upsilon}$ is the Jacobian matrix with elements defined by the partial derivatives of the predicted Stokes parameters S'_j with respect to the free model parameters α_k,

$$\Upsilon^k_j = \frac{\partial S'_j}{\partial \alpha_k}. \tag{I9}$$

In principle, the row index j covers all observations of all 4 Stokes parameters of all source states included as constraints. However, for brevity, the observation index will be ignored and S will loop over the four Stokes parameters, such that j indexes only the source state and $S'_j \in \{I'_j, Q'_j, U'_j, V'_j\}$.

The Hessian is ill-conditioned if there is a high degree of multicollinearity between the unknown model parameters, such that the gradient vector (or column of the Jacobian) for a model parameter is (nearly) proportional to a linear combination of the gradient vectors for a set of other parameters.

To characterize the multicollinearity between model parameters in the approximation of a fixed dipole array, consider a simplified model of the observed Stokes parameters in which the unknown instrumental response consists of only a Lorzentz boost along the Stokes V axis. For the n^{th} source used as a constraint,

$$\underline{\underline{\boldsymbol{S}}}'_n = \underline{\underline{\mathbf{B}}}_V(\beta)\underline{\underline{\mathbf{P}}}\,\underline{\underline{\boldsymbol{S}}}_n. \tag{I10}$$

The set of unknown model parameters include the boost rapidity β and the Stokes parameters intrinsic to each source. Let the model of the intrinsic Stokes parameters for each source be decomposed as

$$\underline{\underline{\boldsymbol{S}}}_n = \underline{\underline{\boldsymbol{S}}}_{n,L} + \underline{\underline{\boldsymbol{S}}}_{n,V}, \tag{I11}$$

where $\underline{\underline{\boldsymbol{S}}}_{n,L} = [I_n, Q_n, U_n, 0]^T$ and $\underline{\underline{\boldsymbol{S}}}_{n,V} = [0,0,0,V_n]^T$, such that

$$\underline{\underline{\boldsymbol{S}}}'_n = \underline{\underline{\mathbf{B}}}_V(\beta)\left(\underline{\underline{\mathbf{P}}}\,\underline{\underline{\boldsymbol{S}}}_{n,L} + \underline{\underline{\boldsymbol{S}}}_{n,V}\right). \tag{I12}$$

With reference to Equation (D7), the partial derivatives of $\underline{\underline{\boldsymbol{S}}}'_n$ with respect to β include only

$$\begin{aligned}\frac{\partial I'_j}{\partial \beta} &= 2\left(I''_j \sinh 2\beta + V_j \cosh 2\beta\right)\\ \frac{\partial V'_j}{\partial \beta} &= 2\left(I''_j \cosh 2\beta + V_j \sinh 2\beta\right),\end{aligned} \tag{I13}$$

where I''_j is the total intensity of the j^{th} source after passing through the projection transformation. Similarly, the partial derivatives of $\underline{\underline{\boldsymbol{S}}}'_j$ with respect to V_n include only

$$\begin{aligned}\frac{\partial I'_j}{\partial V_n} &= \delta_{jn}\sinh 2\beta\\ \frac{\partial V'_j}{\partial V_n} &= \delta_{jn}\cosh 2\beta.\end{aligned} \tag{I14}$$

To first order, the total intensity changes by only a small fraction as the projection transformation varies. For example, even when the differential gain is as large as 2, $\gamma = \ln(2)/2 \sim 0.35$ (Eqn. 69) and $\cosh\gamma \sim 1.06$. Therefore, assume that $I''_n \simeq I_n$ and consider the following linear combination of V_n gradient vectors,

$$\boldsymbol{\Upsilon}_V = \sum_n I_n \boldsymbol{\Upsilon}_{V_n}, \tag{I15}$$

where $\boldsymbol{\Upsilon}_{V_n}$ is the gradient vector defined by the column of the Jacobian matrix corresponding to V_n. The vector $\boldsymbol{\Upsilon}_V$ has components,

$$\Upsilon_{V,j} = \sum_n I_n \frac{\partial S'_j}{\partial V_n} = I_j \frac{\partial S'_j}{\partial V_j} \tag{I16}$$

and its inner product with the gradient vector for β,

$$\begin{aligned} \boldsymbol{\Upsilon}_V \cdot \boldsymbol{\Upsilon}_\beta &= \sum_j I_j \frac{\partial S'_j}{\partial V_j}\frac{\partial S'_j}{\partial \beta}. \\ &\simeq 2\sum_j \left[I_j^2 X(\beta) + I_j V_j Y(\beta)\right] \end{aligned} \tag{I17}$$

where

$$\begin{aligned} X(\beta) &= \sinh^2 2\beta + \cosh^2 2\beta \\ Y(\beta) &= \sinh 2\beta \cosh 2\beta. \end{aligned} \tag{I18}$$

Similarly,

$$|\boldsymbol{\Upsilon}_V|^2 \simeq \sum_j I_j^2 X(\beta) \tag{I19}$$

and

$$|\boldsymbol{\Upsilon}_\beta|^2 \simeq 4\sum_j \left[(I_j^2 + V_j^2)X(\beta) + I_j V_j Y(\beta)\right] \tag{I20}$$

For small β, $X(\beta) \simeq 1$ and $Y(\beta) \simeq 2\beta$. If V_j/I_j is also small, then (to first order) terms involving V_j^2 and βV_j can be ignored, and the cosine similarity between $\boldsymbol{\Upsilon}_V$ and $\boldsymbol{\Upsilon}_\beta$,

$$\frac{\boldsymbol{\Upsilon}_V \cdot \boldsymbol{\Upsilon}_\beta}{|\boldsymbol{\Upsilon}_V||\boldsymbol{\Upsilon}_\beta|} \simeq \frac{2\sum I_j^2}{\left(\sum I_j^2\right)^{\frac{1}{2}}\left(4\sum I_j^2\right)^{\frac{1}{2}}} = 1$$

Owing to the high multicollinearity between V_n and β, the Hessian matrix is poorly-conditioned, its inversion is numerically unstable, and the formal uncertainties of β and V_n are inflated.

REFERENCES

Aab, A., Abreu, P., Aglietta, M., et al. 2014, Phys. Rev. D, 89, 052002, doi: 10.1103/PhysRevD.89.052002

Asad, K. M. B., Koopmans, L. V. E., Jelić, V., et al. 2016, MNRAS, 462, 4482, doi: 10.1093/mnras/stw1863

Bailes, M., Jameson, A., Abbate, F., et al. 2020, PASA, 37, e028, doi: 10.1017/pasa.2020.19

Barakat, R. 1963, J. Opt. Soc. Am., 53, 317

Barakat, R. 1981, Optics Communications, 38, 159, doi: https://doi.org/10.1016/0030-4018(81)90313-8

Beck, R., Brandenburg, A., Moss, D., Shukurov, A., & Sokoloff, D. 1996, ARA&A, 34, 155

Born, M., & Wolf, E. 1980, Principles of optics: electromagnetic theory of propagation, interference and diffraction of light (New York: Pergamon)

Bracewell, R. 1965, The Fourier Transform and its Applications (New York: McGraw–Hill)

Brentjens, M. A., & de Bruyn, A. G. 2005, A&A, 441, 1217, doi: 10.1051/0004-6361:20052990

Britton, M. C. 2000, ApJ, 532, 1240

Broderick, A. E., & Pesce, D. W. 2020, ApJ, 904, 126, doi: 10.3847/1538-4357/abbd9d

Byrne, R. L., Morales, M. F., Hazelton, B., Sullivan, I., & Barry, N. 2022, PASA, 39, e023, doi: 10.1017/pasa.2022.21

Caleb, M., van Straten, W., Keane, E. F., et al. 2019, MNRAS, 487, 1191, doi: 10.1093/mnras/stz1352

Camilo, F., Ransom, S. M., Halpern, J. P., et al. 2006, Nature, 442, 892, doi: 10.1038/nature04986

Cardoso, J.-F. 1991, in ICASSP 91. 1991 International Conference on Acoustics, Speech, and Signal Processing, ed. Y.-T. Chan & A. Venetsanopoulos (New York: IEEE), 3109 – 3112, doi: http://dx.doi.org/10.1109/ICASSP.1991.150113

Clarke, D. 2010, Stellar Polarimetry (New York: Wiley)

Cloude, S. 1986, Optik, 75, 26

Conway, R. G., & Kronberg, P. P. 1969, MNRAS, 142, 11

Cordes, J. M., Shannon, R. M., & Stinebring, D. R. 2016, ApJ, 817, 16, doi: 10.3847/0004-637X/817/1/16

de Gasperin, F., Dijkema, T. J., Drabent, A., et al. 2019, A&A, 622, A5, doi: 10.1051/0004-6361/201833867

Desvignes, G., Kramer, M., Lee, K., et al. 2019, Science, 365, 1013, doi: 10.1126/science.aav7272

Dey, S., Kansabanik, D., Oberoi, D., & Mondal, S. 2025, ApJ, 988, L73, doi: 10.3847/2041-8213/adef0e

Eatough, R. P., Falcke, H., Karuppusamy, R., et al. 2013, Nature, 501, 391, doi: 10.1038/nature12499

Edwards, R. T., & Stappers, B. W. 2004, A&A, 421, 681

Event Horizon Telescope Collaboration, Akiyama, K., Algaba, J. C., et al. 2021, ApJ, 910, L13, doi: 10.3847/2041-8213/abe4de

Fano, U. 1957, Reviews of Modern Physics, 29, 74, doi: 10.1103/RevModPhys.29.74

Ferrière, K., West, J. L., & Jaffe, T. R. 2021, MNRAS, 507, 4968, doi: 10.1093/mnras/stab1641

Frigo, M., & Johnson, S. G. 2005, Proc. IEEE, 93, 216

Ganguly, S., Van Bavel, G. H., & Brown, A. 2000, J. Geophys. Res., 105, 16063, doi: 10.1029/1999JA000238

Goldberg, A. Z., de la Hoz, P., Björk, G., et al. 2021, Advances in Optics and Photonics, 13, 1, doi: 10.1364/AOP.404175

Gould, D. M., & Lyne, A. G. 1998, MNRAS, 301, 235

Granet, C., & James, G. 1997, IEEE Trans. Antennas and Propagation, 45, 1366, doi: 10.1109/8.623125

Grueff, G., Alvito, G., Ambrosini, R., et al. 2004, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5489, Ground-based Telescopes, ed. J. M. Oschmann, Jr., 773–783, doi: 10.1117/12.550332

Guillemot, L., Cognard, I., van Straten, W., Theureau, G., & Gérard, E. 2023, A&A, 678, A79, doi: 10.1051/0004-6361/202347018

Guillemot, L., van Straten, W., Cognard, I., et al. 2025, A&A, 699, A104, doi: 10.1051/0004-6361/202553929

Hamaker, J. P. 2000, A&AS, 143, 515

Hamaker, J. P., & Bregman, J. D. 1996, A&AS, 117, 161

Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137

Han, J. L. 2017, Ann. Rev. Astr. Ap., 55, 111, doi: 10.1146/annurev-astro-091916-055221

Han, J. L., Demorest, P. B., van Straten, W., & Lyne, A. G. 2009, ApJS, 181, 557, doi: 10.1088/0067-0049/181/2/557

Han, J. L., Manchester, R. N., Lyne, A. G., Qiao, G. J., & van Straten, W. 2006, ApJ, 642, 868

Han, J. L., Manchester, R. N., van Straten, W., & Demorest, P. 2018, ApJS, 234, 11, doi: 10.3847/1538-4365/aa9c45

Heiles, C., Perillat, P., Nolan, M., et al. 2001, PASP, 113, 1274

Hotan, A. W., Bailes, M., & Ord, S. M. 2005, ApJ, 624, 906, doi: 10.1086/429270

Hotan, A. W., van Straten, W., & Manchester, R. N. 2004, PASA, 21, 302

Hotan, A. W., Bunton, J. D., Chippendale, A. P., et al. 2021, PASA, 38, e009, doi: 10.1017/pasa.2021.1

IAU. 1974, in Transactions of the IAU, ed. G. Contopoulos & A. Jappel, Vol. XVB (Dordrecht, Netherlands: D. Reidel Pub. Co.), 165 – 167

IEEE. 1983, Std 145-1983, IEEE Standard Definitions of Terms for Antennas (Institute of Electrical and Electronics Engineers)

Islam, M. N., Robishaw, T., Veidt, B., Brar, P., & Du, X. 2024, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 13094, Ground-based and Airborne Telescopes X, ed. H. K. Marshall, J. Spyromilio, & T. Usuda (Bellingham, WA: SPIE), 130943G, doi: 10.1117/12.3020829

Johnston, S. 2002, PASA, 19, 277

Johnston, S., Karastergiou, A., Mitra, D., & Gupta, Y. 2008, MNRAS, 388, 261, doi: 10.1111/j.1365-2966.2008.13379.x

Kansabanik, D., Mondal, S., & Oberoi, D. 2024, ApJ, 968, 55, doi: 10.3847/1538-4357/ad43e9

Kansabanik, D., Vourlidas, A., Dey, S., Mondal, S., & Oberoi, D. 2025, ApJS, 278, 26, doi: 10.3847/1538-4365/adc443

Kawabata, K. 1964, PASJ, 16, 30

Kildal, P. S., Baker, L. A., & Hagfors, T. 1994, Proc. IEEE, 82, 714, doi: 10.1109/5.284738

Lenc, E., Anderson, C. S., Barry, N., et al. 2017, PASA, 34, e040, doi: 10.1017/pasa.2017.36

Liao, Y.-W., Chang, T.-C., Kuo, C.-Y., et al. 2016, ApJ, 833, 289, doi: 10.3847/1538-4357/833/2/289

Lockman, F. J. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 278, Single-Dish Radio Astronomy: Techniques and Applications, ed. S. Stanimirovic, D. Altschuler, P. Goldsmith, & C. Salter, 397–411

Lu, S.-Y., & Chipman, R. A. 1996, J. Opt. Soc. Am. A, 13, 1106, doi: 10.1364/JOSAA.13.001106

Manchester, R. N. 1972, ApJ, 172, 43

Maxwell, J. C. 1865, Philosophical Transactions of the Royal Society of London Series I, 155, 459

McKinnon, M. M. 1992, A&A, 260, 533

Melrose, D. B., & Macquart, J.-P. 1998, ApJ, 505, 921

Mertens, F. G., Mevius, M., Koopmans, L. V. E., et al. 2020, MNRAS, 493, 1662, doi: 10.1093/mnras/staa327

Michilli, D., Seymour, A., Hessels, J. W. T., et al. 2018, Nature, 553, 182, doi: 10.1038/nature25149

Mitchell, D. A., Wayth, R. B., Bernardi, G., Greenhill, L. J., & Ord, S. M. 2012, Journal of Astronomical Instrumentation, 1, 1250003, doi: 10.1142/S2251171712500031

Navarro, J., Manchester, R. N., Sandhu, J. S., Kulkarni, S. R., & Bailes, M. 1997, ApJ, 486, 1019

Noutsos, A., Johnston, S., Kramer, M., & Karastergiou, A. 2008, MNRAS, 386, 1881, doi: 10.1111/j.1365-2966.2008.13188.x

Offringa, A. R., de Bruyn, A. G., Biehl, M., et al. 2010, MNRAS, 405, 155, doi: 10.1111/j.1365-2966.2010.16471.x

Ord, S. M., van Straten, W., Hotan, A. W., & Bailes, M. 2004, MNRAS, 352, 804

Papoulis, A. 1965, Probability, Random Variables, and Stochastic Processes (Sydney: McGraw-Hill)

Perrin, F. 1942, The Journal of Chemical Physics, 10, 415, doi: 10.1063/1.1723743

Petrov, L., Natusch, T., Weston, S., et al. 2015, PASP, 127, 516, doi: 10.1086/681965

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6, doi: 10.1051/0004-6361/201833910

Polnarev, A. G. 1985, Soviet Ast., 29, 607

Porayko, N. K., Noutsos, A., Tiburzi, C., et al. 2019, MNRAS, 483, 4100, doi: 10.1093/mnras/sty3324

Porayko, N. K., Mevius, M., Hernández-Pajares, M., et al. 2023, Journal of Geodesy, 97, 116, doi: 10.1007/s00190-023-01806-1

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1986, Numerical Recipes: The Art of Scientific Computing (Cambridge: Cambridge University Press)

Primak, N., Tiburzi, C., van Straten, W., Dyks, J., & Gulyaev, S. 2022, A&A, 657, A34, doi: 10.1051/0004-6361/202140811

Radhakrishnan, V., & Cooke, D. J. 1969, Astrophys. Lett., 3, 225

Robishaw, T., & Heiles, C. 2021, in The WSPC Handbook of Astronomical Instrumentation, Volume 1: Radio Astronomical Instrumentation, ed. A. Wolszczan (New Jersey: World Scientific), 127–158, doi: 10.1142/9789811203770_0006

Rogers, A. F., van Straten, W., Gulyaev, S., et al. 2024, ApJ, 973, 94, doi: 10.3847/1538-4357/ad656e

Samuel, J., Nityananda, R., & Thyagarajan, N. 2022, Phys. Rev. Lett., 128, 091101, doi: 10.1103/PhysRevLett.128.091101

Schellart, P., Buitink, S., Corstanje, A., et al. 2014, JCAP, 2014, 014, doi: 10.1088/1475-7516/2014/10/014

Serylak, M., Johnston, S., Kramer, M., et al. 2021, MNRAS, 505, 4483, doi: 10.1093/mnras/staa2811

Simon, R. 1982, Optics Communications, 42, 293, doi: https://doi.org/10.1016/0030-4018(82)90234-6

Smirnov, O. M. 2011a, A&A, 527, A106, doi: 10.1051/0004-6361/201016082

—. 2011b, A&A, 527, A107, doi: 10.1051/0004-6361/201116434

—. 2011c, A&A, 527, A108, doi: 10.1051/0004-6361/201116435

—. 2011d, A&A, 531, A159, doi: 10.1051/0004-6361/201116764

Smirnov, O. M., & Tasse, C. 2015, MNRAS, 449, 2668, doi: 10.1093/mnras/stv418

Sokolowski, M., Colegate, T., Sutinjo, A. T., et al. 2017, PASA, 34, e062, doi: 10.1017/pasa.2017.54

Staveley-Smith, L., Wilson, W. E., Bird, T. S., et al. 1996, PASA, 13, 243

Stinebring, D. R., Cordes, J. M., Rankin, J. M., Weisberg, J. M., & Boriakoff, V. 1984, ApJS, 55, 247

Stokes, G. 1852, Trans. Camb. Phil. Soc., 9, 399

Suresh, A., & Cordes, J. M. 2019, ApJ, 870, 29, doi: 10.3847/1538-4357/aaf004

Sutinjo, A., O'Sullivan, J., Lenc, E., et al. 2015, Radio Science, 50, 52, doi: 10.1002/2014RS005517

Trippe, S. 2014, Journal of Korean Astronomical Society, 47, 15, doi: 10.5303/JKAS.2014.47.1.15

Trott, C. M., Jordan, C. H., Murray, S. G., et al. 2018, ApJ, 867, 15, doi: 10.3847/1538-4357/aae314

Van Eck, C. L., Haverkorn, M., Alves, M. I. R., et al. 2017, A&A, 597, A98, doi: 10.1051/0004-6361/201629707

van Straten, W. 2004, ApJS, 152, 129, doi: 10.1086/383187

—. 2013, ApJS, 204, 13, doi: 10.1088/0067-0049/204/1/13

van Straten, W., & Bailes, M. 2011, PASA, 28, 1, doi: 10.1071/AS10021

van Straten, W., Bailes, M., Britton, M., et al. 2001, Nature, 412, 158

van Straten, W., Demorest, P., & Osłowski, S. 2012, Astronomical Research and Technology, 9, 237

van Straten, W., Manchester, R. N., Johnston, S., & Reynolds, J. E. 2010, PASA, 27, 104, doi: 10.1071/AS09084

van Straten, W., & Tiburzi, C. 2017, ApJ, 835, 293, doi: 10.3847/1538-4357/835/2/293

von Hoensbroech, A., & Xilouris, K. M. 1997, A&AS, 126, 121

Wijnholds, S. J., Bregman, J. D., & van Ardenne, A. 2011, Radio Science, 46, doi: https://doi.org/10.1029/2011RS004733

Xilouris, K. M. 1991, A&A, 248, 323

Zheleznyakov, V. V. 1968, Ap&SS, 2, 417, doi: 10.1007/BF02175918