Additive Gaussian Noise in Polarimetric Imaging Systems

Israel J. Vaughn
College of Optical Sciences, University of Arizona, USA
ivaughn@optics.arizona.edu

http://www.optics.arizona.edu/asl

Abstract: Goudail and Bénière claim [10] that in certain circumstances, specifically in a polarimeter with systemic independent additive Gaussian white noise (AGWN), fewer measurements are better [4]. This claim is counter to what is derived in most statistical estimation or imaging science books [2, 6]. In this paper, I analyze conditions in which fewer measurements (or statistical samples) result in better estimation of an object through an imaging system for independent additive Gaussian white noise when using the maximum likelihood estimate. I also seek to derive a better estimator using the full vector electric field propagation through a polarimetric instrument to derive a more robust imaging operator.

OCIS codes: (120.0280) Remote sensing and sensors; (120.5410) Polarimetry; (030.6140) Speckle

---

References and links

1. R. M. A. Azzam. Photopolarimetric measurement of the mueller matrix by fourier analysis of a single detected signal. *Opt. Lett.*, 2(6):148–150, Jun 1978.
2. Harrison H. Barrett and Kyle J. Myers. *Foundations of Image Science*. Wiley-Interscience, Hoboken, New Jersey, United States, 2004.
3. Dennis Goldstein. *Polarized Light*. Marcel Dekker, New York, Basel, 2003.
4. François Goudail and Arnaud Bénière. Estimation precision of the degree of linear polarization and of the angle of polarization in the presence of different sources of noise. *Appl. Opt.*, 49(4):683–693, Feb 2010.
5. S. Huard. *Polarized optical wave*, pages 1–35. Polarization of Light. Wiley, New York, 1997.
6. Steven M. Kay. *Fundamentals of Statistical Signal Processing : Estimation Theory*. Prentice Hall PTR, Upper Saddle River, New Jersey, United States, 1993.
7. P. D. Lax. *Functional Analysis*. Wiley-Interscience, New York, New York, 2002.
8. J. Scott Tyo. Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic error. *Applied Optics*, 41 (4):619–630, 2002.
9. J. Scott Tyo, Charles F. LaCasse, and Bradley M. Ratliff. Total elimination of sampling errors in polarization imagery obtained with integrated microgrid polarimeters. *Opt. Lett.*, 34(20):3187–3189, Oct 2009.
10. Israel J. Vaughn and Brian G. Hoover. Demonstrations of noise and error-reduction algorithms in a rotating-quartz laser polarimeter. volume 6682, page 66820I. SPIE, 2007.
11. Emil Wolf. *Introduction to the Theory of Coherence and the Polarization of Light*. Cambridge University Press, New York, United States, 2007.
12. X. J. Yu and H. S. Kwok. Optical wire-grid polarizers at oblique angles of incidence. *Journal of Applied Physics*, 93(8):4407–4412, 2003.

---

1. Introduction

In this paper, capital bold letters denote matrices, while lower case bold letters denote vectors, unless otherwise stated. It is well known in estimation theory that generally for AGWN, more measurements lower the cost function of the estimation (or observer). If a linear imaging operator $\mathscr{H} : \mathbb{R}^n \to \mathbb{R}^q$ describes the transformation then in general [2]

$$\mathbf{g}(\mathbf{r}) = \mathscr{H}\mathbf{f}(\mathbf{r}) + \mathbf{n}, \quad \mathbf{f} \in \mathbb{R}^{\mathbf{n}},\ \mathbf{g} \in \mathbb{R}^{\mathbf{q}},\ \mathbf{n} \in \mathbb{R}^{\mathbf{q}} \tag{1.1}$$

where $\mathbf{r} \in \mathbb{R}^{\mathbf{3}}$. It is well known that if $\mathscr{H}$ is linear, and $pr(\mathbf{g}|\theta)$ is a (multi-variate) Gaussian distribution, then the Wiener estimator :

$$\hat{\theta} = K_{\overline{\mathbf{g}},\theta} \overline{K}_{\mathbf{g}}^{-1} (\mathbf{g} - \overline{\mathbf{g}}) + \overline{\theta} \tag{1.2}$$

is optimal for the estimated mean squared error (EMSE) cost metric.

2. Polarimeter equations used by current researchers

There are many different types of polarimeters used in practice. They can be active or passive. Both measure the polarization state of the reflected light, but passive polarimeters measure "naturally" lit scenes while active polarimeters control the polarization states of the illuminating light. All current polarimetric instruments depend on two schemes, either measurement of each polarimetric parameter in parallel (which generally requires at least four independent optical paths, and has not been implemented for an active polarimeter) or modulation of some optical parameter which is then used to reconstruct the original optical vector field. The modulation can be in time; examples are dual (or single for passive) rotating retarder (DRR), photoelastic modulators (PEM), dual (or single for passive) rotating polarizers (DRP), etc. The modulation can be in space; examples are microgrid (much like a Bayer color filter), microgrid with retarders, etc. The modulation can be in wavelength; examples are spectropolarimeters.

The polarization of light is a result of the vector nature in 4-dimensional space time and the constraints of Maxwell's equations. Polarimeters are designed to measure polarization in linear media (or reflected from linear media, i.e. the change of polarization properties due to interaction with a sample is linear). The full second order statistical description of polarized light uses a 4×4 mean matrix and a 4×4 covariance matrix between vector fields in each vector element. In optics, 3 mean and covariance matrices are typically used because the time dependence is averaged out. This is because optical detectors cannot typically measure the phase, they actually integrate the vector field over time. The possible spectral coherence effects due to the wave nature of light is described by scalar coherence theory, which can be generalized to the full vector description. To describe a linear transformation between the statistical description (mean and covariance matrices) would require a tensor/Kronecker product formulation since matrices are essentially being transformed to matrices. Due to tensor formulations being somewhat difficult, they are often eschewed in favor of simpler, but inexact, descriptions. Two other mathematical formalisms are often employed in practice, Jones vectors and Stokes vectors. Jones vectors assume that propagation is in the positive z direction in Cartesian coordinates and that the interaction with the media is linear resulting in a 2 element complex valued vector and 2×2 matrices denoting the linear transforms between Jones vectors. This representation is adequate if the light is purely polarized because of the constraint of the wave equation. I will not go into detail on the Jones formalism, see Wolf [11] or Goldstein [3] for an in depth discussion. Stokes vectors are 4 element real valued vectors with 4×4 matrices denoting the linear transforms between Stokes vectors, these matrices are called Mueller matrices. Stokes vectors represent purely and partially polarized states of polarization. Both Jones and Stokes representations do not properly account for interaction between states of polarization due to propagation (even in linear, isotropic, homogeneous media!), see Wolf [11] for details.

Stokes vectors use irradiance measurements of linear combinations of the light projected onto linear and circular states (basically the inner product of the actual light with basis states

Stokes Vector Element	Filter	Description
s_0	I_0	No filter - Total Light, Intensity, or Irradiance
s_1	$I_{\text{HLP}} - I_{\text{VLP}}$	Tendency for Linear Horizontal Polarization
s_2	$I_{\text{45LP}} - I_{\text{135LP}}$	Tendency for Linear 45° Polarization
s_3	$I_{\text{RCP}} - I_{\text{LCP}}$	Tendency for Circular Polarization

Table 1. Stokes parameters using polarizers (analyzers).

using physical polarization analyzer elements). The linear combinations of the projections are *not* unique, resulting in different methods used to measure a single Stokes vector. A commonly used method is the scheme shown in Table 1. I_0 is the irradiance with no polarizer, I_{HLP} is a linear polarizer aligned to transmit light polarized in the horizontal direction, I_{VLP} is a linear polarizer aligned to transmit light polarized in the vertical direction, I_{45LP} is a linear polarizer aligned to transmit light polarized in the 45° to the x-axis direction, I_{135LP} is a linear polarizer aligned to transmit light polarized in the 135° to the x-axis direction, I_{RCP} is a circular polarizer that transmits right circularly polarized light, and I_{LCP} is a circular polarizer that transmits left circularly polarized light. Stokes vectors have some inherent mathematical issues, i.e. the space of physical Stokes vectors is not complete and forms a 4-dimensional cone in $\mathbb{R}^4$ [8], and hence is not a vector space. This can cause issues when attempting to simplify operations by using the linear dual space for inner products. They are, however, the most commonly used formalism in the community.

Generally the DRR types of polarimeters are used in the laboratory setting because of their ease of use and calibration. It is usually assumed that what is being measured is constant over the modulation time, or effectively constant over the modulation time. In a recent (soon to be published) paper by LaCasse *et al* and paper by Tyo *et al* [9], it has been shown that this constraint can be extended through the use of linear systems theory and a band limited assumption on the frequency of the object (either temporal or spatial depending on modulation scheme). In this paper I restrict to the constant case (i.e., no spatially varying or time varying signal on the order of the modulation time).

The equation most practitioners use is linear algebraic and begins with the physical representation :

$$\mathbf{s}_{\text{meas}}(\mathbf{r},t,\lambda) = \mathbf{Y}(\mathbf{r},t,\lambda)\cdot\mathbf{M}\cdot\mathbf{W}(\mathbf{r},t,\lambda)\cdot\mathbf{s}_{\text{in}}(\mathbf{r},t,\lambda) \tag{2.1}$$

where $\mathbf{s}_{\text{in}}$ is the input Stokes vector, $\mathbf{W}$ is the modulation matrix of the incident Stokes vector (usually realized via a polarization element like a polarizer or retarder), $\mathbf{M}$ is the unknown Mueller matrix being measured, $\mathbf{Y}$ is the modulation matrix of the Stokes vector reflected from or transmitted through a sample, and $\mathbf{s}_{\text{meas}}$ is the Stokes vector that exits from the modulating element.

Because the electromagnetic vector field is in the optical frequency range, no currently available detector can directly measure $\mathbf{s}_{\text{meas}}$. If such a detector existed, then $\mathbf{W}$ and $\mathbf{Y}$ would not needed. Optical detectors can only measure some averaged square amplitude of the electric field, which corresponds to the first element of the measured Stokes vector; $s_{0,\text{ meas}}$. Typically what is done is to fix $\mathbf{s}_{\text{in}}$ using some optical device or element, then modulate $\mathbf{W}$ and $\mathbf{Y}$, and finally measure the $s_{0,\text{ meas}}$ to build up a list of values from which the Mueller matrix, $\mathbf{M}$, can be estimated. If $\mathbf{W}$ and $\mathbf{s}_{\text{meas}}$ are known well (through precise engineering and measurements) then it can be assumed that the output Stokes vector from $\mathbf{W}\cdot\mathbf{s}_{\text{in}}$ is known, or has little noise. The noise can also be modeled if desired. Here I use w_{ij} to denote the elements of $\mathbf{W}$ and $y_{i'j'}$ to denote the elements of $\mathbf{Y}$. It is easy to compute that the measured irradiance corresponding to the first element of the Stokes vector is the following from the above matrix equation :

$$
\begin{aligned}
s_{0,\text{ meas}} = \\
& y_{00}(m_{00}\left(w_{00}s_0+w_{01}s_1+w_{02}s_2+w_{03}s_3\right)+m_{01}\left(w_{10}s_0+w_{11}s_1+w_{12}s_2+w_{13}s_3\right) \\
&\quad +m_{02}\left(w_{20}s_0+w_{21}s_1+w_{22}s_2+w_{23}s_3\right)+m_{03}\left(w_{30}s_0+w_{31}s_1+w_{32}s_2+w_{33}s_3\right)) \\
& +y_{01}(m_{10}\left(w_{00}s_0+w_{01}s_1+w_{02}s_2+w_{03}s_3\right)+m_{11}\left(w_{10}s_0+w_{11}s_1+w_{12}s_2+w_{13}s_3\right) \\
&\quad +m_{12}\left(w_{20}s_0+w_{21}s_1+w_{22}s_2+w_{23}s_3\right)+m_{13}\left(w_{30}s_0+w_{31}s_1+w_{32}s_2+w_{33}s_3\right)) \\
& +y_{02}(m_{20}\left(w_{00}s_0+w_{01}s_1+w_{02}s_2+w_{03}s_3\right)+m_{21}\left(w_{10}s_0+w_{11}s_1+w_{12}s_2+w_{13}s_3\right) \\
&\quad +m_{22}\left(w_{20}s_0+w_{21}s_1+w_{22}s_2+w_{23}s_3\right)+m_{23}\left(w_{30}s_0+w_{31}s_1+w_{32}s_2+w_{33}s_3\right)) \\
& +y_{03}(m_{30}\left(w_{00}s_0+w_{01}s_1+w_{02}s_2+w_{03}s_3\right)+m_{31}\left(w_{10}s_0+w_{11}s_1+w_{12}s_2+w_{13}s_3\right) \\
&\quad +m_{32}\left(w_{20}s_0+w_{21}s_1+w_{22}s_2+w_{23}s_3\right)+m_{33}\left(w_{30}s_0+w_{31}s_1+w_{32}s_2+w_{33}s_3\right))
\end{aligned} \tag{2.2}
$$

for an active polarimeter. Please note that in practice the w_{ij} and $y_{i'j'}$ elements are functions of time or space or both, while the input elements, s_k, are fixed. This allows for a solvable system of linear equations. Various measurement schemes can simplify the above general linear equation to obtain **M** from at least 16 or more measurements. This equation can also be analyzed in the Fourier domain, see Azzam [1] for details.

This equation can be represented more compactly as

$$\mathbf{I} = \mathbf{R}\cdot\mathbf{M}\cdot\mathbf{T} \tag{2.3}$$

where **I** is a matrix of modulated $s_{0,\text{ meas}}$ elements (irradiances) from the above equation, **T** is the transmission measurement matrix, of which the rows consist of incident modulated Stokes vectors, and **R** is the received measurement matrix, of which the columns consist of the Stokes vectors exiting the modulating element after the sample being measured. These matrices are not necessarily square.

A passive polarimeter only measures the outgoing Stokes vector instead of the Mueller matrix. For a Stokes (passive) polarimeter Eqn. 2.1 reduces to

$$\mathbf{s}_{\text{meas}}(\mathbf{r},t,\lambda) = \mathbf{W}(\mathbf{r},t,\lambda)\cdot\mathbf{s}_{\text{in}}(\mathbf{r},t,\lambda) \tag{2.4}$$

Which dictates that the Stokes (passive) polarimeter measurement equation reduces to :

$$s_{0,\text{ meas}} = w_{00}s_0+w_{01}s_1+w_{02}s_2+w_{03}s_3 \tag{2.5}$$

for each irradiance measurement where $\mathbf{s}_{\text{in}}$ being estimated. This equation can then be represented more compactly as

$$\mathbf{I} = \mathbf{T}\cdot\mathbf{s}_{\text{in}} \tag{2.6}$$

where **I** is a vector of modulated $s_{0,\text{ meas}}$'s and **T** is the measurement matrix. If noise is included then

$$\mathbf{I} = \mathbf{T}\cdot\mathbf{s}_{\text{in}} + \mathbf{n} \tag{2.7}$$

where **n** is not necessarily linear in addition, i.e. the noise term above can represent additive, multiplicative, or other non-linear or linear noise statistics. Next a recent claim made by Goudail [4] is evaluated.

3. Goudail's Claim

Goudail uses a different, but still common, representation for the Stokes vector :

$$\mathbf{s} = s_0 \begin{pmatrix} 1 \\ P\cos(2\alpha) \\ P\sin(2\alpha) \\ \frac{s_3}{s_0} \end{pmatrix} \tag{3.1}$$

and he ignores the circular retardance term, s_3. Here $P = \frac{\sqrt{s_1^2+s_2^2}}{s_0}$ is the degree of linear polarization (DOLP), and $\alpha \in (-\pi/2, \pi/2]$ is the angle of polarization (AOP) with respect to the reference axis of the measurement system. This is a valid representation, see Huard [5] for details. So Goudail is estimating the linear Stokes vector :

$$\mathbf{s} = s_0 \begin{pmatrix} 1 \\ P\cos(2\alpha) \\ P\sin(2\alpha) \end{pmatrix} \tag{3.2}$$

3.1. Goudail's assumptions

Goudail assumes [4] that to estimate $\mathbf{s}$, the light irradiance is measured at the "output of a set of N linear analyzers are oriented with angles θ_i , $i \in [1, N]$. Depending on the experimental setup, these N measurements may be performed by N different static devices or by the same device that is rotated N times, mechanically or with help of an electro-optic device." I went over the different implementations of a polarimeter in Section 2. Goudail is stating that some modulation scheme is used to obtain N measurements of of the s_0 element, which is then used to estimate $\mathbf{s}$.

The next assumption Goudail makes is important to the claim of using fewer measurements to increase the performance of the estimate. Goudail assumes that there is a constant number of photons incident upon the polarimeter per total measurement [4], i.e. if the system is divided spatially (different optical paths) as described in Section 2 then the incident power is divided equally amongst the optical paths so that each measurement of s_0 is utilizing only $1/N$ photons. This assumption can also be used in other types of polarimeters, each measurement of s_0 only utilizes $1/N$ of some constant irradiance. Since Goudail specifies linear polarizers (analyzers) to measure s_0, the following equation is arrived at using the constant photon assumption

$$s_{0,meas,i} = \frac{1}{2N}\left(s_0 + s_1\cos 2\theta_i + s_2 \sin 2\theta_i\right), \quad 1 \le i \le N \tag{3.3}$$

$$\tag{3.4}$$

$$= \frac{s_0}{2N}\left[1 + P\cos(2\theta_i - 2\alpha)\right] \quad \text{from Eqn. 3.2.} \tag{3.5}$$

The $1/2$ in the above equation comes from the fact that a physical polarizer (analyzer) element projects all incoming light onto one linear polarization state, therefore reducing the power available to the detector by a factor of two. The list of measured irradiances can be represented as

$$\mathbf{I} = \frac{1}{N}\mathbf{T}\cdot\mathbf{s}_{\text{in}} + \mathbf{n} \tag{3.6}$$

where

$$\mathbf{T} = \frac{1}{2}\begin{pmatrix} 1 & \cos 2\theta_1 & \sin 2\theta_1 \\ & \vdots & \\ 1 & \cos 2\theta_i & \sin 2\theta_i \\ & \vdots & \\ 1 & \cos 2\theta_N & \sin 2\theta_N \end{pmatrix} \tag{3.7}$$

and $\mathbf{n}$ is noise. Goudail states that the pseudoinverse (also known as linear least squares) of $\mathbf{T}$ is the optimal estimator, in the maximum likelihood (ML) sense, of $\mathbf{s}_{\text{in}}$ when the noise is additive, white and Gaussian distributed (AWGN). Hence, the estimate of $\mathbf{s}_{\text{in}}$ given Eqn. 3.6 is the following :

$$\hat{\mathbf{s}}_{\text{in}} = N\mathbf{T}^{+}\mathbf{I} \tag{3.8}$$

where the $+$ denotes pseudoinverse and

$$\mathbf{T}^{+} = (\mathbf{T}^{t}\mathbf{T})^{-1}\mathbf{T}^{t} \tag{3.9}$$

where $\mathbf{T}^{t}$ is the transpose of $\mathbf{T}$. If $\langle \mathbf{n} \rangle = 0$ then the estimator above is unbiased, where $\langle \cdot \rangle$ denotes the ensemble mean.

Goudail then assumes that the noise vector, $\mathbf{n}$ is independent in the statistical sense and that it is also independent of the irradiance. Goudail then states that the covariance matrix of the estimator is

$$\mathbf{K}_{\hat{\mathbf{s}}_{\text{in}}} = N^2\mathbf{T}^{+}\mathbf{K}_{\mathbf{I}}(\mathbf{T}^{+})^{t} \tag{3.10}$$

where $\mathbf{K}_{\mathbf{I}}$ is the covariance matrix of $\mathbf{I}$. Since $\mathbf{n}$ is AWGN

$$\mathbf{K}_{\mathbf{I}} = \sigma^2\mathbf{1}_N \tag{3.11}$$

where $\mathbf{1}_N$ is the $N \times N$ identity matrix. When this is substituted into Eqn. 3.10

$$\begin{aligned} \mathbf{K}_{\hat{\mathbf{s}}_{\text{in}}} &= N^2\sigma^2\mathbf{T}^{+}\mathbf{1}_N(\mathbf{T}^{+})^{t} \\ &= N^2\sigma^2(\mathbf{T}^{t}\mathbf{T})^{-1}\mathbf{T}^{t}\left[(\mathbf{T}^{t}\mathbf{T})^{-1}\mathbf{T}^{t}\right]^{t} \\ &= N^2\sigma^2(\mathbf{T}^{t}\mathbf{T})^{-1}\mathbf{T}^{t}\mathbf{T}[(\mathbf{T}^{t}\mathbf{T})^{-1}]^{t} \\ &= N^2\sigma^2(\mathbf{T}^{t}\mathbf{T})^{-1} \end{aligned} \tag{3.12}$$

Goudail states "The covariance matrix is a function of the measurement matrix $\mathbf{T}$, which depends on the chosen measurement angles θ_i. This choice constitutes a degree of freedom of the measurement system, which can be used to optimize its performance. A classical performance criterion is the trace of $\mathbf{K}_{\hat{\mathbf{s}}_{\text{in}}}$, which represents the sum of the variances on the three components of the Stokes vector." He then goes on to show that equally spaced angles minimize the trace. Goudail then uses these assumptions to derive a result that states that fewer measurements increase the performance.

3.2. Fewer measurements yield better performance

Using equally spaced angles on $(-\pi/2, \pi/2]$ it can be easily shown that

$$(\mathbf{T}^{t}\mathbf{T})^{-1} = \frac{4}{N}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{3.13}$$

since

$$\sum_{i=1}^{N}\cos 2\theta_i = \sum_{i=1}^{N}\sin 2\theta_i = \sum_{i=1}^{N}\cos 2\theta_i \sin 2\theta_i = 0, \tag{3.14}$$

$$\sum_{i=1}^{N}\cos^2 2\theta_i = \sum_{i=1}^{N}\sin^2 2\theta_i = \frac{N}{2}. \tag{3.15}$$

Substituting this into Eqn. 3.9 yields

$$\mathbf{T}^+ = \frac{2}{N}\begin{pmatrix} 1 & \cdots & 1 \\ 2\cos 2\theta_1 & \cdots & 2\cos 2\theta_N \\ 2\sin 2\theta_1 & \cdots & 2\sin 2\theta_N \end{pmatrix}. \tag{3.16}$$

Finally Eqn. 3.13 can be substituted into Eqn. 3.12 to get

$$\boxed{\mathbf{K}_{\hat{\mathbf{s}}_{\text{in}}} = 4N\sigma^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}} \tag{3.17}$$

Goudail then goes on to state that if the noise is independent of the number of measurements and irradiance, then fewer measurements will result in lower magnitudes in the elements of the covariance matrix shown in Eqn. 3.17. This is strictly true if his system analysis is correct, and the variance of the noise is not dependent on the irradiance.

4. Maximum likelihood estimation in the presence of AWGN

Goudail assumes in his claim that following equation describes the imaging system:

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n} \tag{4.1}$$

where $\mathbf{n}$ is independent (element to element) and Gaussian distributed. Eqn. 3.12 is general for this case so

$$\mathbf{K}_{\hat{\mathbf{f}}} = \sigma^2(\mathbf{H}^t\mathbf{H})^{-1} \tag{4.2}$$

for the pseudoinverse estimator $\mathbf{H}^+$. It is well known that the pseudoinverse estimator is equivalent to the optimal (linear least squares) maximum likelihood estimator for the independent Gaussian noise case [6]. Here $\mathbf{f}$ is $1 \times N$ and $\mathbf{g}$ is $1 \times M$ which forces $\mathbf{H}$ to be $M \times N$. It must also be assumed that $M \geq N$, i.e. an overdetermined system. If $\mathbf{H}$ is represented in the following form

$$\mathbf{H} = \begin{pmatrix} \mathbf{h}_1^t \\ \vdots \\ \mathbf{h}_M^t \end{pmatrix} \tag{4.3}$$

where

$$\mathbf{h}_j = \begin{pmatrix} h_{j1} \\ \vdots \\ h_{jN} \end{pmatrix} \tag{4.4}$$

then

$$(\mathbf{H}^t\mathbf{H}) = \begin{pmatrix} \mathbf{h}_1\cdot\mathbf{h}_1 & \cdots & \mathbf{h}_1\cdot\mathbf{h}_M \\ \mathbf{h}_2\cdot\mathbf{h}_1 & \cdots & \mathbf{h}_2\cdot\mathbf{h}_M \\ \vdots & \ddots & \vdots \\ \mathbf{h}_M\cdot\mathbf{h}_1 & \cdots & \mathbf{h}_M\cdot\mathbf{h}_M \end{pmatrix} \tag{4.5}$$

where $\mathbf{h}_j\cdot\mathbf{h}_k$ is the inner (dot) product. It is obvious that $(\mathbf{H}^t\mathbf{H})$ is symmetric. This matrix is always invertible by construction. If the covariance matrix $\mathbf{K}_{\hat{\mathbf{f}}}$ is some scalar function of the number of measurements, M, say $p(M) > 0$ for all M, i.e.

$$\mathbf{K}_{\hat{\mathbf{f}}} = \sigma^2(\mathbf{H}^t\mathbf{H})^{-1} = \sigma^2 p(M)(\widetilde{\mathbf{H}}^t\widetilde{\mathbf{H}})^{-1} \tag{4.6}$$

then

$$(\mathbf{H}^t\mathbf{H}) = \frac{1}{p(M)}(\widetilde{\mathbf{H}}^t\widetilde{\mathbf{H}}). \tag{4.7}$$

This forces each $\mathbf{h}_j$ to be exactly

$$\mathbf{h}_j = \frac{1}{\sqrt{p(M)}}\widetilde{\mathbf{h}}_j \tag{4.8}$$

which forces

$$\mathbf{H} = \frac{1}{\sqrt{p(M)}}\widetilde{\mathbf{H}} \tag{4.9}$$

If the above equation was not true, then Eqn. 4.6 would be violated for some particular value of M.

The $\mathbf{h}_j$s correspond (for a polarimeter) to measurement, they could also correspond to pixels,etc. This is an interesting case, where each individual measurement row depends on the total number of measurements taken. In imaging science $\mathbf{H}$ is fixed. In this case, if M is fixed, then the normal predictions of estimation theory hold, i.e. the estimation of $\mathbf{f}$ will improve if multiple $\mathbf{g}$s for the same $\mathbf{f}$ are measured, basically a "meta-measurement" for Goudail. What Goudail is really doing, in imaging terms, is optimizing $\mathbf{H}$, the system parameter, which is a different process. In imaging, the noise analysis is done using a fixed $\mathbf{H}$, which does not preclude optimizing the system parameters.

A relevant question is "What are the statistics of $\mathbf{g}$ if $\mathbf{H}$ is allowed to vary with M?" Eqn. 4.1 can be re-written as

$$\mathbf{g} = \frac{1}{\sqrt{p(M)}}\widetilde{\mathbf{H}}\mathbf{f} + \mathbf{n}. \tag{4.10}$$

This implies that the mean of $\mathbf{g}$ is

$$\langle\mathbf{g}\rangle = \left\langle \frac{1}{\sqrt{p(M)}}\widetilde{\mathbf{H}}\mathbf{f} \right\rangle + \langle\mathbf{n}\rangle \tag{4.11}$$

$$= \frac{1}{\sqrt{p(M)}}\widetilde{\mathbf{H}}\mathbf{f} + \boldsymbol{\mu}_n \tag{4.12}$$

if $\mathbf{f}$ is not random and $\mathbf{n}$ is normally distributed with mean $\boldsymbol{\mu}_n$ and covariance $\mathbf{K}_n$. Here

$$\boldsymbol{\mu}_n = \begin{pmatrix} \mu_n \\ \mu_n \\ \vdots \\ \mu_n \end{pmatrix} = \mu_n \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{4.13}$$

and

$$\mathbf{K}_n = \begin{pmatrix} \sigma_n^2 & 0 & \cdots & 0 \\ 0 & \sigma_n^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{pmatrix} = \sigma_n^2 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \sigma_n^2 \mathbf{1}. \tag{4.14}$$

In this case each component of $\mathbf{g}$ has mean

$$\langle g_j \rangle = \frac{1}{\sqrt{p(M)}} \widetilde{\mathbf{h}}_j \cdot \mathbf{f} + \mu_n \tag{4.15}$$

which is dependent on M. Each component of the covariance matrix is

$$\begin{aligned} \mathrm{Cov}(g_j, g_k) &= \Big\langle \big(g_j - \langle g_j \rangle\big)\big(g_k - \langle g_k \rangle\big) \Big\rangle && (4.16) \\ &= \left\langle \left(g_j - \frac{1}{\sqrt{p(M)}} \widetilde{\mathbf{h}}_j \cdot \mathbf{f} - \mu_n \right) \left(g_k - \frac{1}{\sqrt{p(M)}} \widetilde{\mathbf{h}}_k \cdot \mathbf{f} - \mu_n \right) \right\rangle && (4.17) \\ &= \big\langle (n_j - \mu_n)(n_k - \mu_n) \big\rangle && (4.18) \\ &= K_{n,jk} && (4.19) \end{aligned}$$

So yes, for a fixed M, $\mathbf{g}$ is a Gaussian distributed random variable. The covariance matrix is also independent of M. But if M is a free parameter, then the mean of $\mathbf{g}$ is dependent on M, the distribution on $\mathbf{g}$ changes with M, and linear least squares is no longer necessarily optimal. This should be investigated further, but I am out of time.

5. Passive polarimeter equation and optimal estimator for the AWGN case

Note : I began this calculation to attempt to find a better estimator than the linear one for a polarimeter. Due to time constraints this section was unable to be completed, but is still relevant to a discussion of how to measure the electric vector field.

In this section I investigate the Stokes polarimeter equation in an imaging sense, and attempt to derive the optimal estimator when $\mathbf{n}$ is AWGN.

In general the following equation describes an imaging system,

$$\mathbf{g}(\mathbf{r},t) = \mathscr{H}\mathbf{f}(\mathbf{r},t) + \mathbf{n}, \quad f \in \mathbb{R}^4,\ g \in \mathbb{R}^4,\ n \in \mathbb{R}^4, \tag{5.1}$$

where $\mathbf{f}(\mathbf{r},t)$ is the full vector electric field of electromagnetic radiation exiting our object, $\mathscr{H}$ describes an operator that changes the vector field in some way, and $\mathbf{g}(\mathbf{r},t)$ is the full vector field after the operator. $\mathscr{H}$ corresponds to a series of polarization elements, while the exiting $\mathbf{g}(\mathbf{r},t)$ will be measured on a detector which collapses the electric vector field to a scalar irradiance measurement. This must be done multiple times, changing $\mathscr{H}$ for each measurement in order to reconstruct the original vector field.

5.1. Poynting vector and plane waves

Any vector field whose components are elements of the Schwartz class can be represented as a possibly infinite but countable superposition of plane waves [7], so **f** can represented as

$$\mathbf{g}(\mathbf{r},t) = \begin{pmatrix} \sum_{n_1} E_{n_1} \exp\left[i\left(\mathbf{k}_{n_1}\cdot\mathbf{r} - \omega_{n_1} t\right)\right] \\ \sum_{n_2} E_{n_2} \exp\left[i\left(\mathbf{k}_{n_2}\cdot\mathbf{r} - \omega_{n_2} t\right)\right] \\ \sum_{n_3} E_{n_3} \exp\left[i\left(\mathbf{k}_{n_3}\cdot\mathbf{r} - \omega_{n_3} t\right)\right] \end{pmatrix} \tag{5.2}$$

$$= \sum_j \mathbf{e}_j \exp\left[i\left(\mathbf{k}_j\cdot\mathbf{r} - \omega_j t\right)\right] \tag{5.3}$$

where the $E_{n_1}, E_{n_2}, E_{n_3}$ are possibly complex field amplitudes and $\mathbf{e}_j = (E_{n_1}, E_{n_2}, E_{n_2})$ for some specific n_1, n_2, n_3, i.e. j is the re-indexing of n_1, n_2, n_3. The sum can be reordered since it is countable and uniformly convergent (if it wasn't then an arbitrary field couldn't be represented as a superposition of plane waves, this restricts the forms the sum can take mathematically). In this specific case, it can be assumed that **f** is statistically stationary in time, given constant illumination of an object. For example, in a sunlight scene, typically the sun position with respect to the observed object is moving very slowly compared with acquisition time, and the light illuminating the object being imaged has undergone multiple scattering events, which can be modeled using stochastic processes. If **n** is AWGN, then it too is statistically stationary, which implies that **g** is stationary. To model the measurement of **g** the Poynting vector is used;

$$\mathbf{s}(\mathbf{r},t) = \mathrm{Re}\left[\mathbf{e}(\mathbf{r},t)\right] \times \mathrm{Re}\left[\mathbf{h}(\mathbf{r},t)\right] \tag{5.4}$$

where $\times$ denotes the cross product, **e** is the complex electric field, and **h** is the complex magnetic (induction) field. It can be shown that the time average of a Poynting vector for an electromagnetic plane wave in a homogeneous, isotropic, linear medium, i.e.

$$\begin{aligned} \text{Electric vector field} &= \mathrm{Re}[\mathbf{e}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}] \\ \text{Magnetic vector field} &= \mathrm{Re}[\mathbf{h}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}], \end{aligned} \tag{5.5}$$

where $\mathbf{e}_0 = \mathbf{e}_0^{'} + i\mathbf{e}_0^{''}$ is the possibly complex electric amplitude, $\mathbf{h}_0 = \mathbf{h}_0^{'} + i\mathbf{h}_0^{''}$ is the possibly complex magnetic amplitude, and $\mathbf{k} = \mathbf{k}^{'} + i\mathbf{k}^{''}$ is the possibly complex wave vector; can be represented in the following form :

$$\langle \mathbf{s}(\mathbf{r})\rangle_t = \frac{e^{-2\mathbf{k}''\cdot\mathbf{r}}}{2} \frac{\left(\|\mathbf{e}_0^{'}\|^2 + \|\mathbf{e}_0^{''}\|^2\right)(\mu'\mathbf{k}' + \mu''\mathbf{k}'') - 2(\mu'\mathbf{k}'' - \mu''\mathbf{k}') \times (\mathbf{e}_0^{'} \times \mathbf{e}_0^{''})}{z_0(\omega/c)(\mu'^2 + \mu''^2)}. \tag{5.6}$$

$\mu(\omega) = \mu^{'}(\omega) + i\mu^{''}(\omega) = 1 + \chi_m(\omega)$ is the complex magnetic permeability and z_0 is the impedance of free space. This assumes that $\mathbf{s}(\mathbf{r},t)$ is statistically stationary. Note that the $\times$ symbols in the above equation denote cross products. If the plane wave is in a typical optical medium then $\mathrm{Im}[\varepsilon(\omega)] \approx 0$ and $\mathrm{Im}[\mu(\omega)] \approx 0$. Additionally if $\mathbf{k}'' = 0$ then Eqn. 5.6 can be simplified to

$$\langle \mathbf{s}(\mathbf{r})\rangle_t = \frac{1}{2} \frac{\left(\|\mathbf{e}_0^{'}\|^2 + \|\mathbf{e}_0^{''}\|^2\right)}{z_0\sqrt{\mu(\omega)/\varepsilon(\omega)}} \hat{\mathbf{k}} \tag{5.7}$$

where $\hat{\mathbf{k}} = \mathbf{k}/\|\mathbf{k}\|$. The irradiance at a single point, **r**, is usually defined as the time average of the Poynting vector.

Equation 5.7 is for a single plane wave, so I must generalize the equation for an arbitrary superposition of plane waves. For an arbitrary superposition of plane waves

$$\mathbf{s}(\mathbf{r},t) = \mathrm{Re}\left(\sum_j \mathbf{e}_j \exp\left[i\left(\mathbf{k}_j \cdot \mathbf{r} - \omega_j t\right)\right]\right) \times \mathrm{Re}\left(\sum_j \mathbf{h}_j \exp\left[i\left(\mathbf{k}_j \cdot \mathbf{r} - \omega_j t\right)\right]\right). \tag{5.8}$$

Since

$$\mathrm{Re}(z) = \frac{1}{2}\left(z + z^*\right) \tag{5.9}$$

where $*$ denotes the complex conjugate Eqn. 5.8 can be rewritten

$$\mathbf{s}(\mathbf{r},t) = \frac{1}{4}\sum_j \sum_{j'} \begin{pmatrix} \mathbf{e}_j \times \mathbf{h}_{j'} \exp\left[i(\mathbf{k}_j + \mathbf{k}_{j'}) \cdot \mathbf{r} - i(\omega_j + \omega_{j'})t\right] \\ + \quad \mathbf{e}_j \times \mathbf{h}^*_{j'} \exp\left[i(\mathbf{k}_j - \mathbf{k}_{j'}) \cdot \mathbf{r} - i(\omega_j - \omega_{j'})t\right] \\ + \quad \mathbf{e}^*_j \times \mathbf{h}_{j'} \exp\left[-i(\mathbf{k}_j - \mathbf{k}_{j'}) \cdot \mathbf{r} + i(\omega_j - \omega_{j'})t\right] \\ + \quad \mathbf{e}^*_j \times \mathbf{h}^*_{j'} \exp\left[-i(\mathbf{k}_j + \mathbf{k}_{j'}) \cdot \mathbf{r} + i(\omega_j + \omega_{j'})t\right] \end{pmatrix} \tag{5.10}$$

In the above equation the terms inside the parentheses are added because the equation will not fit on a single line, the parentheses do not represent a vector. When $j = j'$ the time average of $\mathbf{s}$ for that portion of the sum is the same as in Eqn. 5.6. What about when $j \neq j'$? Then the electric field is not associated with the magnetic field, and hence the cross product is not perpendicular to either plane wave. The time average can be taken inside the sum and then the following four integrals must be evaluated :

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[-i(\omega_j + \omega_{j'})t\right] dt \tag{5.11}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[-i(\omega_j - \omega_{j'})t\right] dt \tag{5.12}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[i(\omega_j - \omega_{j'})t\right] dt \tag{5.13}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[i(\omega_j + \omega_{j'})t\right] dt \tag{5.14}$$

Since they are sinusoids, they average to zero in all except the following special cases :

$$\omega_j = \omega_{j'}, \tag{5.15}$$

$$\omega_j = -\omega_{j'}. \tag{5.16}$$

From Maxwell's equations, each plane wave component has the following property

$$\mathbf{k}_j \cdot \mathbf{k}_j = \mathbf{k}'_j \cdot \mathbf{k}'_j - \mathbf{k}''_j \cdot \mathbf{k}''_j + 2i\mathbf{k}'_j \cdot \mathbf{k}''_j = \omega_j^2 \varepsilon_0 \mu_0 \varepsilon(\omega_j)\mu(\omega_j) = \left(\frac{\omega_j}{c}\right)^2 \varepsilon(\omega_j)\mu(\omega_j) \tag{5.17}$$

but after the polarizer elements which constitute $\mathscr{H}$ the field is usually propagating in air, which is close enough to free space such that $\varepsilon(\omega_j) = \mu(\omega_j) \approx 1$. This implies there is no imaginary component of the equation so

$$\mathbf{k}_j \cdot \mathbf{k}_j = \mathbf{k}'_j \cdot \mathbf{k}'_j = \left(\frac{\omega_j}{c}\right)^2 \tag{5.18}$$

If $\omega_j = \omega_{j'}$ then

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[-i(\omega_j + \omega_{j'})t\right] dt = 0 \tag{5.19}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[-i(\omega_j - \omega_{j'})t\right] dt = 1 \tag{5.20}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[i(\omega_j - \omega_{j'})t\right] dt = 1 \tag{5.21}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[i(\omega_j + \omega_{j'})t\right] dt = 0 \tag{5.22}$$

and if $\omega_j = -\omega_{j'}$ then

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[-i(\omega_j + \omega_{j'})t\right] dt = 1 \tag{5.23}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[-i(\omega_j - \omega_{j'})t\right] dt = 0 \tag{5.24}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[i(\omega_j - \omega_{j'})t\right] dt = 0 \tag{5.25}$$

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \exp\left[i(\omega_j + \omega_{j'})t\right] dt = 1. \tag{5.26}$$

This implies that

$$\langle \mathbf{s}(\mathbf{r},t)\rangle_t = \frac{1}{4}\sum_j \sum_{j'} \begin{pmatrix} & \mathbf{e}_j \times \mathbf{h}_{j'} \exp\left(i(\mathbf{k}_j + \mathbf{k}_{j'})\cdot\mathbf{r}\right) \\ + & \mathbf{e}_j \times \mathbf{h}^*_{j'} \exp\left(i(\mathbf{k}_j - \mathbf{k}_{j'})\cdot\mathbf{r}\right) \\ + & \mathbf{e}^*_j \times \mathbf{h}_{j'} \exp\left(-i(\mathbf{k}_j - \mathbf{k}_{j'})\cdot\mathbf{r}\right) \\ + & \mathbf{e}^*_j \times \mathbf{h}^*_{j'} \exp\left(-i(\mathbf{k}_j + \mathbf{k}_{j'})\cdot\mathbf{r}\right) \end{pmatrix} \tag{5.27}$$

where the j' indices are restricted to only those where $\omega_j = \omega_{j'}$ or $\omega_j = -\omega_{j'}$. Eqn. 5.18 also forces the following condition :

$$\|\mathbf{k}_j\|^2 = \|\mathbf{k}_{j'}\|^2 \tag{5.28}$$

This means that any corresponding $\mathbf{k}_{j'}$ vector must lie on the sphere of radius $\|\mathbf{k}_j\|$.

5.2. Polarization components and the vector field

The polarization change can be explicitly computed using material properties, but $\mathscr{H}$ is used instead to describe the change of the vector field through a polarization component. In the complex plane wave representation that is being used, the real and imaginary parts of $\mathbf{e}_j$, along with the real parts of the exponential, define the polarization properties. Explicitly

$$\mathbf{e}(\mathbf{r},t) = \mathrm{Re}\left[\mathbf{e}_0 e^{\mathbf{k}\cdot\mathbf{r} - \omega t}\right] \tag{5.29}$$

$$= e^{-\mathbf{k}''\cdot\mathbf{r}}\,\mathrm{Re}\left[(\mathbf{e}'_0 + i\mathbf{e}''_0)e^{i(\mathbf{k}'\cdot\mathbf{r} - \omega t)}\right] \tag{5.30}$$

$$= e^{-\mathbf{k}''\cdot\mathbf{r}}\left[\mathbf{e}'_0 \cos(\mathbf{k}'\cdot\mathbf{r} - \omega t) - \mathbf{e}''_0 \sin(\mathbf{k}'\cdot\mathbf{r} - \omega t)\right] \tag{5.31}$$

Since the plane waves before the polarizer elements are propagating in air, $\mathbf{k}'' \approx 0$. So $\mathscr{H}$ acts on

$$\boxed{\mathbf{e}(\mathbf{r},t) = \mathbf{e}'_0 \cos(\mathbf{k}'\cdot\mathbf{r} - \omega t) - \mathbf{e}''_0 \sin(\mathbf{k}'\cdot\mathbf{r} - \omega t)} \tag{5.32}$$

Ideally, the polarizer transmission does not depend on the amplitude terms (the sine and cosine terms) in Eqn. 5.32, so only the vector terms matter. The polarizer must be oriented in space, so the coordinates can be chosen. For this analysis, it is assumed that the polarizer can be modeled as a plane element, with normal pointing in the $\hat{\mathbf{z}}$ direction. In reality, wire grid polarizers work well for a range of angles less than $\pm\pi/2$ with respect to $\hat{\mathbf{z}}$. The physics change as the angle becomes large due to Brewster's angle (of the substrate) and other effects [12], these details will not be included here. This analysis assumes that the polarizer is ideal over the entire range of angles $(-\pi/2, \pi/2]$ and that it is large with respect to the angular subtense of the electric field being measured. Since only the plane waves propagating from left to right can pass through the polarizer, any plane waves propagating from right to left are omitted. The polarizer will be defined to have the transmission axis at angle θ with respect to the positive x-axis. If the polarizer it thin, just like the wires, and/or on a dielectric substrate, then only reflection of the plane waves affect the transmission of the z components, this is partially responsible for the polarizer performance as a function of angle with respect to $\hat{\mathbf{z}}$.

The components of each plane wave must first be projected onto the xy-plane :

$$\mathbf{e}'_{\text{proj}} = \mathbf{e}' - (\mathbf{e}' \cdot \hat{\mathbf{z}})\hat{\mathbf{z}} \tag{5.33}$$

$$\mathbf{e}''_{\text{proj}} = \mathbf{e}'' - (\mathbf{e}'' \cdot \hat{\mathbf{z}})\hat{\mathbf{z}} \tag{5.34}$$

i.e. first each component is projected onto the normal, $\hat{\mathbf{z}}$, then subtracted that projection to get the component on the xy-plane. Then only the components oriented in the θ direction are transmitted, i.e.

$$\mathbf{e}'_{\text{trans,xy}} = \left[\mathbf{e}'_{\text{proj}} \cdot \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}\right] \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \tag{5.35}$$

$$\mathbf{e}''_{\text{trans,xy}} = \left[\mathbf{e}''_{\text{proj}} \cdot \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}\right] \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \tag{5.36}$$

The above determines the x, y components of both $\mathbf{e}', \mathbf{e}''$, with the z components remaining unchanged. Simplifying the above results in

$$\mathbf{e}'_{\text{trans}} = \begin{pmatrix} (e'_x \cos\theta + e'_y \sin\theta)\cos\theta \\ (e'_x \cos\theta + e'_y \sin\theta)\sin\theta \\ e'_z \end{pmatrix} \tag{5.37}$$

$$\mathbf{e}''_{\text{trans}} = \begin{pmatrix} (e''_x \cos\theta + e''_y \sin\theta)\cos\theta \\ (e''_x \cos\theta + e''_y \sin\theta)\sin\theta \\ e''_z \end{pmatrix} \tag{5.38}$$

for a single plane wave. As a consequence of the superposition principle the total field is

$$\mathbf{e}_{\text{trans}} = \sum_j \left[\begin{pmatrix} (e'_{x,j}\cos\theta + e'_{y,j}\sin\theta)\cos\theta \\ (e'_{x,j}\cos\theta + e'_{y,j}\sin\theta)\sin\theta \\ e'_{z,j} \end{pmatrix} \cos(\mathbf{k}'_j\cdot\mathbf{r} - \omega_j t) \right.$$

$$\left. - \begin{pmatrix} (e''_{x,j}\cos\theta + e''_{y,j}\sin\theta)\cos\theta \\ (e''_{x,j}\cos\theta + e''_{y,j}\sin\theta)\sin\theta \\ e''_{z,j} \end{pmatrix} \sin(\mathbf{k}'_j\cdot\mathbf{r} - \omega_j t) \right] \tag{5.39}$$

$$= \sum_j \mathrm{Re}\left\{ \left[\begin{pmatrix} (e'_{x,j}\cos\theta + e'_{y,j}\sin\theta)\cos\theta \\ (e'_{x,j}\cos\theta + e'_{y,j}\sin\theta)\sin\theta \\ e'_{z,j} \end{pmatrix} + i \begin{pmatrix} (e''_{x,j}\cos\theta + e''_{y,j}\sin\theta)\cos\theta \\ (e''_{x,j}\cos\theta + e''_{y,j}\sin\theta)\sin\theta \\ e''_{z,j} \end{pmatrix} \right] e^{\mathbf{k}'_j\cdot\mathbf{r} - \omega_j t} \right\}. \tag{5.40}$$

This implies that

$$\mathscr{H}(\theta) = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta & 0 \\ \cos\theta\sin\theta & \sin^2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.41}$$

Notice that $\mathscr{H}$ is self-adjoint. The propagating electric field is still a superposition of plane waves in basically free space, so from Maxwell's equations the $\mathbf{h}_j$ vectors are mutually perpendicular to both $\mathbf{e}_j$ and $\mathbf{k}'_j$. From Maxwell's equations it can be easily shown that

$$\mathbf{h} = \frac{1}{\mu_0\omega}\mathbf{k}'\times\mathbf{e} \tag{5.42}$$

for each plane wave. The irradiance measured on the detector from Eqns. 5.27, 5.42 is then

$$\langle\mathbf{s}(\mathbf{r},t)\rangle_t = \frac{1}{4}\sum_j\sum_{j'}\frac{1}{\mu_0\omega_{j'}} \begin{pmatrix} \left[(\mathscr{H}\mathbf{e}_j\cdot\mathscr{H}\mathbf{e}_{j'})\mathbf{k}_{j'} - (\mathscr{H}\mathbf{e}_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}_{j'}\right]\exp\left(i(\mathbf{k}_j+\mathbf{k}_{j'})\cdot\mathbf{r}\right) \\ + \left[(\mathscr{H}\mathbf{e}_j\cdot\mathscr{H}\mathbf{e}^*_{j'})\mathbf{k}_{j'} - (\mathscr{H}\mathbf{e}_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}^*_{j'}\right]\exp\left(i(\mathbf{k}_j-\mathbf{k}_{j'})\cdot\mathbf{r}\right) \\ + \left[(\mathscr{H}\mathbf{e}^*_j\cdot\mathscr{H}\mathbf{e}_{j'})\mathbf{k}_{j'} - (\mathscr{H}\mathbf{e}^*_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}_{j'}\right]\exp\left(-i(\mathbf{k}_j-\mathbf{k}_{j'})\cdot\mathbf{r}\right) \\ + \left[(\mathscr{H}\mathbf{e}^*_j\cdot\mathscr{H}\mathbf{e}^*_{j'})\mathbf{k}_{j'} - (\mathscr{H}\mathbf{e}^*_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}^*_{j'}\right]\exp\left(-i(\mathbf{k}_j+\mathbf{k}_{j'})\cdot\mathbf{r}\right) \end{pmatrix} \tag{5.43}$$

by using the vector identity

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}. \tag{5.44}$$

Note that since the plane waves are in air, $\mathbf{k}_j = \mathbf{k}'_j$ and $\mathbf{k}_{j'} = \mathbf{k}'_{j'}$, i.e. all $\mathbf{k}$ vectors are purely real.

Now each component from Eqn. 5.40 can be substituted into Eqn. 5.43. First the inner products in front of the $\mathbf{k}$ vector terms are computed.

$$(\mathscr{H}\mathbf{e}_j\cdot\mathscr{H}\mathbf{e}_{j'})\mathbf{k}_{j'} = [(\mathscr{H}^2\mathbf{e}_j)\cdot\mathbf{e}_{j'}]\mathbf{k}_{j'}, \text{ since } \mathscr{H} \text{ is self-adjoint} \tag{5.45}$$

where

$$\mathscr{H}^2(\theta) = \begin{pmatrix} \cos^4\theta + \cos^2\theta\sin^2\theta & \cos^3\theta\sin\theta + \cos\theta\sin^3\theta & 0 \\ \cos^3\theta\sin\theta + \cos\theta\sin^3\theta & \sin^4\theta + \cos^2\theta\sin^2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.46}$$

Additionally

$$\begin{aligned}(\mathscr{H}^2\mathbf{e}_j \cdot \mathbf{e}_{j'})\mathbf{k}_{j'} &= [(\mathscr{H}^2\mathbf{e}'_j + i\mathscr{H}^2\mathbf{e}'') \cdot (\mathbf{e}'_{j'} + i\mathbf{e}''_{j'})]\mathbf{k}_{j'} \\ &= [\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}'_{j'} + i\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'} + i\mathscr{H}^2\mathbf{e}''_j \cdot \mathbf{e}'_{j'} - \mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'}]\mathbf{k}_{j'}.\end{aligned} \tag{5.47}$$

Similarly

$$(\mathscr{H}^2\mathbf{e}_j \cdot \mathbf{e}^*_{j'})\mathbf{k}_{j'} = [\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}'_{j'} + i\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'} - i\mathscr{H}^2\mathbf{e}''_j \cdot \mathbf{e}'_{j'} + \mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'}]\mathbf{k}_{j'} \tag{5.48}$$

$$(\mathscr{H}^2\mathbf{e}^*_j \cdot \mathbf{e}_{j'})\mathbf{k}_{j'} = [\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}'_{j'} - i\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'} + i\mathscr{H}^2\mathbf{e}''_j \cdot \mathbf{e}'_{j'} + \mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'}]\mathbf{k}_{j'} \tag{5.49}$$

$$(\mathscr{H}^2\mathbf{e}^*_j \cdot \mathbf{e}^*_{j'})\mathbf{k}_{j'} = [\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}'_{j'} - i\mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'} - i\mathscr{H}^2\mathbf{e}''_j \cdot \mathbf{e}'_{j'} - \mathscr{H}^2\mathbf{e}'_j \cdot \mathbf{e}''_{j'}]\mathbf{k}_{j'} \tag{5.50}$$

and explicitly

$$\begin{aligned}(\mathscr{H}^2\mathbf{e}_j \cdot \mathbf{e}_{j'})\mathbf{k}_{j'} = \Big[& e'_{x,j}e'_{x,j'}\cos^4\theta + (e'_{x,j}e'_{y,j'} + e'_{x,j'}e'_{y,j})\cos^3\theta\sin\theta + e'_{y,j}e'_{y,j'}\cos^2\theta\sin^2\theta \\ & + e'_{x,j}e'_{x,j'}\cos^2\theta\sin^2\theta + (e'_{x,j}e'_{y,j'} + e'_{x,j'}e'_{y,j})\sin^3\theta\cos\theta + e'_{y,j}e'_{y,j'}\sin^4\theta \\ & + e'_{z,j}e'_{y,j'} \\ & - e''_{x,j}e''_{x,j'}\cos^4\theta - (e''_{x,j}e''_{y,j'} - e''_{x,j'}e''_{y,j})\cos^3\theta\sin\theta + e''_{y,j}e''_{y,j'}\cos^2\theta\sin^2\theta \\ & - e''_{x,j}e''_{x,j'}\cos^2\theta\sin^2\theta - (e''_{x,j}e''_{y,j'} - e''_{x,j'}e''_{y,j})\sin^3\theta\cos\theta - e''_{y,j}e''_{y,j'}\sin^4\theta \\ & - e''_{z,j}e''_{y,j'} \\ & + i\Big(e''_{x,j}e'_{x,j'}\cos^4\theta + (e''_{x,j}e'_{y,j'} + e'_{x,j'}e''_{y,j})\cos^3\theta\sin\theta + e''_{y,j}e'_{y,j'}\cos^2\theta\sin^2\theta \\ & + e''_{x,j}e'_{x,j'}\cos^2\theta\sin^2\theta + (e''_{x,j}e'_{y,j'} + e'_{x,j'}e''_{y,j})\sin^3\theta\cos\theta + e''_{y,j}e'_{y,j'}\sin^4\theta \\ & + e''_{z,j}e'_{y,j'} \\ & + e'_{x,j}e''_{x,j'}\cos^4\theta + (e'_{x,j}e''_{y,j'} + e''_{x,j'}e'_{y,j})\cos^3\theta\sin\theta + e'_{y,j}e''_{y,j'}\cos^2\theta\sin^2\theta \\ & + e'_{x,j}e''_{x,j'}\cos^2\theta\sin^2\theta + (e'_{x,j}e''_{y,j'} + e''_{x,j'}e'_{y,j})\sin^3\theta\cos\theta + e'_{y,j}e''_{y,j'}\sin^4\theta \\ & + e'_{z,j}e''_{y,j'}\Big)\Big].\end{aligned} \tag{5.51}$$

With similar equations for the other compoenents. The above equation is tedious to manipulate. The detector not only integrates over time, but also over some finite area. If the polarizer is set at $z = 0$ and centered on $(0,0)$ in the xy-plane then the detector is some fixed distance z_0 from the polarizer also centered on $(0,0)$ in the xy-plane. The detector will be assumed to be square here, with side length l. Then

$$\begin{aligned}\text{irradiance on detector} &= \int_{-l/2}^{l/2}\int_{-l/2}^{l/2} \langle \mathbf{s}(\mathbf{r},t)\rangle_t\, dxdy \qquad (5.52) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \operatorname{rect}\left(\frac{x}{l},\frac{y}{l}\right)\langle \mathbf{s}(\mathbf{r},t)\rangle_t\, dxdy. \qquad (5.53)\end{aligned}$$

This integral can be taken inside Eqn. 5.43 and only operates on the exponential terms. The following four integrals are obtained :

$$e^{i(k_{z,j}+k_{z,j'})z_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\text{rect}\left(\frac{x}{l},\frac{y}{l}\right)\exp\left[i(k_{x,j}+k_{x,j'})x+i(k_{y,j}+k_{y,j'})y\right]dxdy \tag{5.54}$$

$$e^{i(k_{z,j}-k_{z,j'})z_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\text{rect}\left(\frac{x}{l},\frac{y}{l}\right)\exp\left[i(k_{x,j}-k_{x,j'})x+i(k_{y,j}-k_{y,j'})y\right]dxdy \tag{5.55}$$

$$e^{-i(k_{z,j}-k_{z,j'})z_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\text{rect}\left(\frac{x}{l},\frac{y}{l}\right)\exp\left[-i(k_{x,j}-k_{x,j'})x-i(k_{y,j}-k_{y,j'})y\right]dxdy \tag{5.56}$$

$$e^{-i(k_{z,j}+k_{z,j'})z_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\text{rect}\left(\frac{x}{l},\frac{y}{l}\right)\exp\left[-i(k_{x,j}+k_{x,j'})x-i(k_{y,j}+k_{y,j'})y\right]dxdy \tag{5.57}$$

These all resemble Fourier transforms. These integrals respectively become

$$l^2 e^{i(k_{z,j}+k_{z,j'})z_0}\,\text{sinc}\left(\frac{l(k_{x,j}+k_{x,j'})}{2\pi},\frac{l(k_{y,j}+k_{y,j'})}{2\pi}\right) \tag{5.58}$$

$$l^2 e^{i(k_{z,j}-k_{z,j'})z_0}\,\text{sinc}\left(\frac{l(k_{x,j}-k_{x,j'})}{2\pi},\frac{l(k_{y,j}-k_{y,j'})}{2\pi}\right) \tag{5.59}$$

$$l^2 e^{-i(k_{z,j}-k_{z,j'})z_0}\,\text{sinc}\left(\frac{l(-k_{x,j}+k_{x,j'})}{2\pi},\frac{l(-k_{y,j}+k_{y,j'})}{2\pi}\right) \tag{5.60}$$

$$l^2 e^{-i(k_{z,j}+k_{z,j'})z_0}\,\text{sinc}\left(-\frac{l(k_{x,j}+k_{x,j'})}{2\pi},-\frac{l(k_{y,j}+k_{y,j'})}{2\pi}\right). \tag{5.61}$$

Finally, the vector field irradiance that can possibly be measured by the detector is :

$$\int_{-l/2}^{l/2}\int_{-l/2}^{l/2}\langle\mathbf{s}(\mathbf{r},t)\rangle_t dxdy =$$

$$\boxed{\frac{1}{4}\sum_j\sum_{j'}\frac{l^2}{\mu_0\omega_{j'}}\begin{pmatrix} & \left[(\mathscr{H}\mathbf{e}_j\cdot\mathscr{H}\mathbf{e}_{j'})\mathbf{k}_{j'}-(\mathscr{H}\mathbf{e}_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}_{j'}\right]e^{i(k_{z,j}+k_{z,j'})z_0}\,\text{sinc}\left(\frac{l(\boldsymbol{\kappa}_{x,j}+\boldsymbol{\kappa}_{x,j'})}{2\pi}\right) \\ + & \left[(\mathscr{H}\mathbf{e}_j\cdot\mathscr{H}\mathbf{e}^*_{j'})\mathbf{k}_{j'}-(\mathscr{H}\mathbf{e}_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}^*_{j'}\right]e^{i(k_{z,j}-k_{z,j'})z_0}\,\text{sinc}\left(\frac{l(\boldsymbol{\kappa}_{x,j}-\boldsymbol{\kappa}_{x,j'})}{2\pi}\right) \\ + & \left[(\mathscr{H}\mathbf{e}^*_j\cdot\mathscr{H}\mathbf{e}_{j'})\mathbf{k}_{j'}-(\mathscr{H}\mathbf{e}^*_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}_{j'}\right]e^{-i(k_{z,j}-k_{z,j'})z_0}\,\text{sinc}\left(\frac{l(-\boldsymbol{\kappa}_{x,j}+\boldsymbol{\kappa}_{x,j'})}{2\pi}\right) \\ + & \left[(\mathscr{H}\mathbf{e}^*_j\cdot\mathscr{H}\mathbf{e}^*_{j'})\mathbf{k}_{j'}-(\mathscr{H}\mathbf{e}^*_j\cdot\mathbf{k}_{j'})\mathscr{H}\mathbf{e}^*_{j'}\right]e^{-i(k_{z,j}+k_{z,j'})z_0}\,\text{sinc}\left(-\frac{l(\boldsymbol{\kappa}_{x,j}+\boldsymbol{\kappa}_{x,j'})}{2\pi}\right)\end{pmatrix}} \tag{5.62}$$

where

$$\boldsymbol{\kappa}=\begin{pmatrix}k_x\\k_y\end{pmatrix} \tag{5.63}$$

$$\text{sinc}(\boldsymbol{\kappa})=sinc(k_x,k_y)=\text{sinc}(k_x)\,\text{sinc}(k_y). \tag{5.64}$$

This must, of course, be projected onto the detector. Typically detectors are flat semiconductor devices, so this can be modeled with a plane at z_0. This implies that Eqn. 5.62 must be projected onto the plane at z_0.

6. Conclusions

Goudail [4] showed that when using the pseudoinverse for polarimetric measurements there exist circumstances where more measurements can result in larger variances in the estimate of a Stokes vector. I showed that this is not true in the imaging science sense, Goudail is optimizing the system operator. I also showed that if the mathematics are reformulated to account for changing the system operator, then the assumptions of a Gaussian $\mathbf{g}$ is incorrect. This analysis should be further investigated in the future to analyze the imaging system Goudail is utilizing in more detail and perhaps find an optimal estimator for the case where the number of measurements is changing.

I also began an analysis to find a more robust polarimetric imaging operator, using the full vector electric field representation and Maxwell's equations. This analysis is the beginning of a potentially fruitful re-formulation of the polarimetric imaging operator.