5-2017

Electrodynamical Modeling for Light Transport Simulation

Michael G. Saunders
East Tennessee State University

Electrodynamical Modeling for Light Transport Simulation

A thesis submitted to the Department of Engineering Technology, Surveying, and Digital Media at East Tennessee State University in partial fulfillment of the requirements for the degree of Bachelor of Science with honors.

Thesis Mentor: Prof. Martin Fitzgerald, M.M.

Thesis Reader: Prof. Jonathan Hounshell, M.F.A.

Thesis Reader: Prof. Richard Ignace, Ph.D.

April 2017

by
Michael Grady Saunders

mgradysaunders@gmail.com

To my mother, father, sister, and lately departed four-legged friend:

Thanks for inspiring my brazen curiosity, and for supporting my endeavors at every opportunity.

— Grady

Abstract

Modernity in the computer graphics community is characterized by a burgeoning interest in physically based rendering techniques. That is to say that mathematical reasoning from first principles is widely preferred to ad hoc, approximate reasoning in blind pursuit of photorealism. Thereby, the purpose of our research is to investigate the efficacy of explicit electrodynamical modeling by means of the generalized Jones vector given by Azzam [1] and the generalized Jones matrix given by Ortega-Quijano & Arce-Diego [2] in the context of stochastic light transport simulation for computer graphics. To augment the status quo path tracing framework with such a modeling technique would permit a plethora of complex optical effects—including dispersion, birefringence, dichroism, and thin film interference, and the physical optical elements associated with these effects—to become naturally supported, fully integrated features in physically based rendering software.

Contents

1. Introduction

"There is the disadvantange of not knowing all languages," said Conseil, "or the disadvantage of not having one universal language."

—Jules Verne, *Twenty Thousand Leagues Under the Sea*

Author's Note

I am well aware of the flock of anxious, unanswered questions that imbues this document with a sense of fickle uncertainty—it is but the same flock that pesters all academic work, I suppose. *Who is the target audience? What are readers expected to be familiar with? Are any noteworthy conventions or symbols employed?* Fear not. I will drive off the proverbial nuisance at once.

Broadly speaking, scientific writings have two goals. Perhaps the more essential is to dignify a result that has not yet been established. The other, which is often under-emphasized, is to educate readers about the ideas that inform the result. Mathematics is unique in that there exist such things as provably correct results, and it is for this reason that mathematically-inclined papers tend to emphasize proofs. Yet, proofs are far from the most palatable sources of knowledge. Informal discourse that aims to develop intuition has been more valuable in my experience than rigorous discourse that abandons analogies, examples, and layman's terms to pursue scrupulous, infallible truth.

So, I maintain a vested interest in making this document accessible to students, tinkerers, and hobbyists even though I aim it at the day-to-day practitioners of computer graphics—academics, professionals, and the like. Notwithstanding, I could not hope to review adequately the material which enables the contributions presented herein. I thusly presume that the reader is well-versed in the language of linear algebra, basic vector calculus, and probability theory. The interested reader should not be discouraged from pressing onward if he or she is presently deficient in any of these areas, however. When I commit to learning something new, I begin by floundering through some reading material that is terribly over my head, all the while doing my best to identify what I do not know and why it matters—it is a good way to get started!

1.1. Light, Matter, and Graphics

Photorealistic rendering intends to fabricate imagery which cannot be distinguished from authentic photography. Although rendering need not be photorealistic, the pursuit of photorealism is a well-established and persisting objective in computer graphics. Lately it has become commonplace to use "photorealistic" and "physically-based" interchangably, but it is important to notice that these terms are not strictly equivalent. That is to say that the emphasis of photorealism is placed on the observer—and specious, untenable methods are fair game! On the other hand, *physically-based* rendering emphasizes logical coherence, mathematical rigor, and the established laws of physics under the presumption that a plausible simulation of the interactions of light and matter will produce photorealistic imagery. And, given the sudden ubiquity of physically-based techniques in recent years, it seems that this presumption holds in practice.

Be that as it may, *time* has proven to be the recurring adversary of graphics applications. Real-time rendering software, of course, is expected to maximize fidelity while maintaining the industry-standard sixty frames per second. De facto time constraints of offline renderers are less demanding (on the order of hours to days), but offline image quality is held to an elevated and ever-increasing standard. With good reason, then, the graphics community is relentlessly pragmatic. Enduring topics of interest include, but are not limited to, acceleration structures, convergence rates, denoising, antialiasing, and error analysis. But the noble pursuit of pragmatism generally leads to a preoccupation with *geometric optics*, the “cheapest” optical theory, which expatriates electromagnetism from the literature in turn.

Geometric optics, also known as ray optics, is laden with geometrical postulates that describe the large-scale behavior of light under ordinary conditions. First and foremost, light is thought to propagate outward from points on semi-infinite lines called *rays*. Among the other postulates are the laws of reflection and refraction, as well as Fermat’s general notion that light takes the path of least time. Although these postulates are justified by more tortuous optical theory, they inherently limit the modeling capabilities of geometric optics. Effects owed to wave and polarization optics are ignored, and they are certainly non-negligible! Structural coloration, which is due to wave interference rather than pigment, is observed in peacock feathers, butterfly wings, soap bubbles, and a variety of other places. More so, the refractive indices of birefringent materials, such as crystals and plastics, as well as the absorption coefficients of dichroic materials, such as dichroic glass, depend on polarization state.

1.2. Overview

We investigate the efficacy of applied wave and polarization optics in the context of light transport simulation for computer graphics, building from the work of Azzam [1] and Ortega-Quijano & Arce-Diego [2] which has established a generalized form of the Jones calculus in the optics literature. To be specific, we review the features of the generalized Jones calculus with emphasis on implementation in physically based rendering software—we identify preliminary concerns posed by the nature of the modeling technique, as well as methods for circumventing them. In addition, we derive a method for computing the generalized Jones matrix of [2] by formulating the matrix exponential as an eigenvalue problem, for which an example implementation is provided.

1.2.1. Conventions

Matrix and vector quantities are typeset in boldface italic. Furthermore, vector quantities associated with planewaves, which appear in equations 1, 3, and 6, are accented with an overbar, and subsequent planewave generalizations are superscripted with an additional prime (the $'$ symbol). Refer to table 1 for the glyphs associated with greek letters.

Physical measurements follow the International System of Units based on meters, kilograms, and seconds—the seven base units are displayed in table 2, and a handful of relevant derived units are displayed in table 3. Electromagnetic constants are consistently presumed to be for the vacuum, so subscripted naughts are left off (for example, vacuum permittivity and vacuum permeability are denoted ε and μ rather than ε_0 and μ_0). In this way, electromagnetic formulas are derived from the microscopic form of Maxwell’s equations, and we deal with neither the electric displacement field $\boldsymbol{D}$ nor the auxillary magnetic field $\boldsymbol{H}$.

1.2.2. Matrix Exponentials

Recall that the exponential function e^x for $x \in \mathbb{C}$ may be expressed as a power series,

$$\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \implies 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

which converges for all $x \in \mathbb{C}$. In section 3, we begin dealing with *matrix exponentials* instead of scalar exponentials. Do not fret, for the exponential of a matrix is strikingly similar to the exponential of a scalar—the exponential function $\mathrm{e}^{\boldsymbol{X}}$ for $\boldsymbol{X} \in \mathbb{C}^{n,n}$ may also be expressed as a power series,

$$\mathrm{e}^{\boldsymbol{X}} = \sum_{k=0}^{\infty} \frac{\boldsymbol{X}^k}{k!} \implies \boldsymbol{I} + \boldsymbol{X} + \frac{\boldsymbol{X}^2}{2!} + \frac{\boldsymbol{X}^3}{3!} + \cdots,$$

which indeed converges for all $\boldsymbol{X} \in \mathbb{C}^{n,n}$. In the same way that the integral powers of a scalar are found by repeated scalar multiplication, the integral powers of a matrix are found by repeated matrix multiplication. Specifically,

$$k = 0 \implies \boldsymbol{X}^0 = \boldsymbol{I},$$
$$k > 0 \implies \boldsymbol{X}^k = \prod_{n=1}^{k} \boldsymbol{X},$$

and, if $\boldsymbol{X}$ is invertible,

$$k < 0 \implies \boldsymbol{X}^k = \prod_{n=1}^{-k} \boldsymbol{X}^{-1}.$$

Because matrices may be multiplied if and only if the number of columns in the first is equal to the number of rows in the second, a matrix must be square to be multiplied by itself. It is for this reason that the exponential function is defined only for square matrices.

Perhaps it goes without saying, but the interpretation of matrix exponentials is non-trivial. For the intents and purposes at hand, it should be sufficient to consider the differential behavior of matrix exponentials. As in the ordinary case,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{tx}\right] = x\mathrm{e}^{tx},$$

the derivative of a matrix exponential with respect to a scalar multiplier t is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{t\boldsymbol{X}}\right] = \boldsymbol{X}\mathrm{e}^{t\boldsymbol{X}}.$$

1.2.3. Matrix Exponentials and Commutativity

It is well known that the exponential of a sum is equal to the product of the exponentials of the terms when the terms are scalars. So, it is important to note that this is *not* necessarily true when the terms are matrices. This is stated concisely in the language of implications,

$$\begin{aligned} a, b \in \mathbb{C}^{1,1} &\implies \mathrm{e}^{a+b} = \mathrm{e}^{a}\mathrm{e}^{b}, \text{ but} \\ \boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n,n} &\not\Longrightarrow \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}}\mathrm{e}^{\boldsymbol{B}}. \end{aligned}$$

The following result is stated without proof. In general, the exponential of a sum is equivalent to the product of the exponentials of the terms if and only if the terms commute under multiplication—that is, multiplication of the terms must not depend on order of appearance. To illustrate, let $\boldsymbol{X}_i \in \mathbb{C}^{n,n}$ denote the i'th of m matrices. We have

$$\mathrm{e}^{\boldsymbol{X}_1+\cdots+\boldsymbol{X}_m} = \mathrm{e}^{\boldsymbol{X}_1} \cdots \mathrm{e}^{\boldsymbol{X}_m} \iff (\forall i, j \in [1, m])(\boldsymbol{X}_i\boldsymbol{X}_j = \boldsymbol{X}_j\boldsymbol{X}_i).$$

Since matrix multiplication is not necessarily commutative, it is often desirable to establish so-called *commutation relations* between matrices. To accomplish this, we employ the *commutator*,

$$[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{A}\boldsymbol{B} - \boldsymbol{B}\boldsymbol{A}.$$

Thereby, two matrices commute if and only if their commutator is the zero matrix.

Uppercase	Lowercase	Letter name
A	α	alpha
B	β	beta
Γ	γ	gamma
Δ	δ	delta
E	ε, ϵ	epsilon
Z	ζ	zeta
H	η	eta
Θ	θ, ϑ	theta
I	ι	iota
K	κ	kappa
Λ	λ	lambda
M	μ	mu
N	ν	nu
Ξ	ξ	xi
O	o	omicron
Π	π, ϖ	pi
P	ρ, ϱ	rho
Σ	σ, ς	sigma
T	τ	tau
Υ	υ	upsilon
Φ	ϕ, φ	phi
X	χ	chi
Ψ	ψ	psi
Ω	ω	omega

Table 1: Greek typeface reference.

Unit			Measure
ampere	(A)	$\longleftrightarrow$	electric current
candela	(cd)	$\longleftrightarrow$	luminous intensity
meter	(m)	$\longleftrightarrow$	length
kilogram	(kg)	$\longleftrightarrow$	mass
second	(s)	$\longleftrightarrow$	time
kelvin	(K)	$\longleftrightarrow$	temperature
mole	(mol)	$\longleftrightarrow$	amount of substance

Table 2: SI base unit reference.

Unit			Measure	Equivalent base units
newton	(N)	$\longleftrightarrow$	force/weight	$kg \cdot m / s^2$
hertz	(Hz)	$\longleftrightarrow$	frequency	$1 / s$
coulomb	(C)	$\longleftrightarrow$	electrical charge	$s \cdot A$
joule	(J)	$\longleftrightarrow$	work/energy	$kg \cdot m^2 / s^2$
henry	(H)	$\longleftrightarrow$	electrical inductance	$kg \cdot m^2 / s^2 / A^2$
farad	(F)	$\longleftrightarrow$	electrical capacitance	$s^4 \cdot A^2 / kg / m^2$
ohm	(Ω)	$\longleftrightarrow$	electrical resistance	$kg \cdot m^2 / s^3 / A^2$
volt	(V)	$\longleftrightarrow$	electrical potential difference	$kg \cdot m^2 / s^3 / A$
watt	(W)	$\longleftrightarrow$	power	$kg \cdot m^2 / s^3$
radian	(rad)	$\longleftrightarrow$	angle	1 *(dimensionless)*
steradian	(sr)	$\longleftrightarrow$	solid angle	1 *(dimensionless)*

Table 3: SI derived unit reference. This table features only those units deemed relevant to the discussion here. There are many more derived units with special names, see [4].

2. The Generalized Jones Vector

> A tune is an example of a non-uniform object. We have perceived it as a whole in a certain duration; but the tune as a tune is not at any moment of that duration though one of the individual notes may be located there.
>
> —Alfred North Whitehead, *The Concept of Nature*

2.1. Getting Started

Before I blather on about polarization as a mathematical concept, I shall first remind you that polarization is an actual phenomenon that actually exists. Moreover, the effects of polarization are readily observed in person, unlike, for instance, the implications of miniature black holes or extra dimensions which, if experimentally suggested at all, will only be suggested by the smouldering mounds of data rolling out of particle accelerators like the Large Hadron Collider in Geneva.

At the time of writing, three dimensional cinema is made possible by the means of polarized glasses. If you happen to have a pair of these glasses lying around, I encourage you to disassemble them in order to look through both lenses in sequence—no disassembly is required if you have two pairs of course. And if you have no such pairs, just take my word for the time being and try this out for yourself in the future. In any case, upon looking through both lenses in sequence you will find a particular angle for each lense at which you can see through to the other side. At that point, however, rotating one lens or the other by ninety degrees to the left or right will cause the lens farthest away from you to appear black. That is to say that rotating one lens or the other by ninety degrees prohibits any light from making it through both lenses. Therefore, light must have some intrinsic directional property by which it may be filtered, which suggests the idea of polarization state.

Yet, we have put the cart before the horse—it is difficult to discuss polarization state any further without embracing the posits of classical electromagnetism and the canonical (not generalized) Jones vector. Let us do so immediately.

2.2. Classical Electromagnetism

Classical electromagnetic theory springs forth from the idea that, at every point in space and time,[1] the forces exterted by electromagnetic phenomena are modeled accurately by a pair of three-vectors satisfying Maxwell's equations:

$$\begin{aligned}
&\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon}, && \textit{(Gauss's law)} \\
&\nabla \cdot \boldsymbol{B} = 0, && \textit{(Gauss's law for magnetism)} \\
&\nabla \times \boldsymbol{E} = -\frac{\partial}{\partial t}\boldsymbol{B}, \text{ and} && \textit{(Faraday's law of induction)} \\
&\nabla \times \boldsymbol{B} = \mu \boldsymbol{J} + \frac{1}{c^2}\frac{\partial}{\partial t}\boldsymbol{E} && \textit{(Ampère's circuital law with Maxwell's addition)}
\end{aligned}$$

[1] Intentionally avoiding the term "spacetime" as not to provoke the beast that is relativity. Albeit Maxwell's equations are consistent with special relativity, the discussion here remains non-relativistic.

where $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{J}$, and ρ denote the electric field vector, the magnetic field vector, the current density vector, and the charge density, and

$$\begin{aligned} c &\approx 2.99 \times 10^8 \text{ m / s}, \\ \varepsilon &\approx 8.85 \times 10^{-12} \text{ F / m, and} \\ \mu &= 4\pi \times 10^{-7} \text{ H / m} \end{aligned}$$

are physical constants indicating the speed of light in free space, the permittivity of free space, and the permeability of free space. It is worth noting that the values of c and ε are known *exactly*, despite the approximate values shown above. This follows from the observation that, in SI units, c is exact by definition, like μ, which determines ε in turn through the identity $c^2\mu\varepsilon = 1$.

Then, when we say that something is an electromagnetic phenomenon, we mean that it is modeled accurately as a disturbance in the classical electromagnetic field. Such disturbances are more often called *waves* in accordance with their general behavior. This in turn is why we say that light, being an electromagnetic phenomenon, propagates as a wave. However, it is important to be specific about the nature of these waves—to liken disturbances in the electromagnetic field to disturbances on the surface of a pond without justification, for instance, can be notably misleading.

Let us imagine that at each and every point there exists an arrow $\boldsymbol{E}$ indicating the strength of the electric field and an arrow $\boldsymbol{B}$ indicating the strength of the magnetic field. Upon introducing a disturbance at a point $\boldsymbol{P}$, the arrows at $\boldsymbol{P}$ change magnitude and direction to reflect the new strengths of the electric and magnetic fields, and Maxwell's equations spark a chain reaction—the arrows nearby must also change magnitude and direction, but then the arrows near those arrows must change, and so on and so forth. This causes the disturbance to propagate outward through space over time and thusly to be intepreted as a wave.

In mathematical terms, the electromagnetic wave equations are formulated by taking the curl of Maxwell's curl equations (i.e., Faraday's law of induction and Ampère's circuital law with Maxwell's addition). Illustrating the first step,

$$\begin{aligned} \nabla \times (\nabla \times \boldsymbol{E}) &= \nabla \times \left(-\frac{\partial}{\partial t}\boldsymbol{B}\right) \text{ and} \\ \nabla \times (\nabla \times \boldsymbol{B}) &= \nabla \times \left(\mu\boldsymbol{J} + \frac{1}{c^2}\frac{\partial}{\partial t}\boldsymbol{E}\right). \end{aligned}$$

To derive the wave equations for the vacuum, we assume there are no charges and no currents so ρ and $\boldsymbol{J}$ vanish. Applying the identity $\nabla \times (\nabla \times \boldsymbol{f}) = \nabla(\nabla \cdot \boldsymbol{f}) - \nabla^2\boldsymbol{f}$, distributing the remaining curl, and resubstituting Maxwell's equations, we obtain

$$\begin{aligned} \nabla^2\boldsymbol{E} &= \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{E} \text{ and} \\ \nabla^2\boldsymbol{B} &= \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\boldsymbol{B} \end{aligned}$$

which indeed signify wave-like behavior. That is, each establishes a linear relation between a second order topological derivative, ∇^2, and a second order temporal derivative, $\partial^2/\partial t^2$. Note that the wave equations for the vacuum may be (and often are) written succinctly as

$$\begin{aligned} \Box^2\boldsymbol{E} &= \boldsymbol{0} \text{ and} \\ \Box^2\boldsymbol{B} &= \boldsymbol{0}, \end{aligned}$$

where $\Box^2$ is the d'Alembertian operator,

$$\Box^2 = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}.$$

Although arbitrarily complicated solutions to the wave equations exist, we are gladly preoccupied by a particularly useful, surprisingly simple solution: *the monochromatic plane wave*. A monochromatic plane wave is a wave (a function of position and time that satsifies the wave equation) whose wavefronts are infinitely many parallel planes characterized by a fixed wavelength λ. In the context of classical electromagnetism, the electric field of a monochromatic plane wave which propagates in free space along the $\hat{\boldsymbol{z}}$ axis with amplitude A and phase offset α is given by

$$\bar{\boldsymbol{E}}\,(z,t) = A\mathrm{e}^{\mathrm{i}(kz-\omega t+\alpha)}\,\bar{\boldsymbol{E}}_0\,(\chi,\psi) \tag{1}$$

where $k = 2\pi/\lambda$ denotes the angular wavenumber, $\omega = 2c\pi/\lambda$ denotes the angular frequency, and $\bar{\boldsymbol{E}}_0$ denotes the *canonical Jones vector* which is parameterized by real angles χ and ψ that determine the polarization mode of the wave. In particular,

$$\bar{\boldsymbol{E}}_0\,(\chi,\psi) = \begin{pmatrix} \cos\chi\cos\psi + \mathrm{i}\,\sin\chi\,\sin\psi \\ \cos\chi\,\sin\psi - \mathrm{i}\,\sin\chi\cos\psi \\ 0 \end{pmatrix}. \tag{2}$$

It is important to notice that $\bar{\boldsymbol{E}}_0$ is a complex unit vector in the xy plane whose real and imaginary parts correspond to the major and minor axes of the ellipse determined by ellipticity angle χ and orientation ψ, as per figure 1. This ellipse is therefore called the *canonical polarization ellipse*, and it is used to classify the fundamental polarization modes. Furthermore, equation 1 traces this ellipse (subject to the scaling factor A) when evaluated over t at a fixed z, or over z at a fixed t.

Referring again to figure 1, it is quite obvious that the polarization ellipse collapses into a line as χ approaches 0. Hence, when $\chi = 0$ we say that the wave is *linearly polarized*. Further distinctions are often drawn based on the orientation ψ. When $\psi = 0$ we may call the wave *horizontally* linearly polarized. Similiarly, when $\psi = \pi/2$ we may call the wave *vertically* linearly polarized. When $\chi \neq 0$ we say in general that the wave is *elliptically polarized*. However, the ellipse forms a perfect circle when $|\chi| = \pi/4$, in which case we specify that the wave is *circularly polarized*. Elliptically polarized light may be qualified further by examining the sign of χ. When χ is positive the electric field vector appears to spin anti-clockwise as it traces out the ellipse. When χ is negative the electric field vector appears to spin clockwise. To communicate which, we may specify that the wave is *left* or *right* elliptically or circularly polarized.

What about the magnetic field?

It may appear that we have only established a model for the electric field vector of a monochromatic plane wave, propagating in free space along the $\hat{\boldsymbol{z}}$ axis. However, the electric field vector precisely determines the magnetic field vector of such a wave via Maxwell's equations. It happens that the equation of the magnetic field vector $\bar{\boldsymbol{B}}$ must be given by

$$\bar{\boldsymbol{B}}(z,t) = \frac{\hat{\boldsymbol{z}} \times \bar{\boldsymbol{E}}(z,t)}{c}.$$

This implies indeed that the electric field vector $\bar{\boldsymbol{E}}$, the magnetic field vector $\bar{\boldsymbol{B}}$, and the propagation direction $\hat{\boldsymbol{z}}$ are mutually perpendicular in this case.

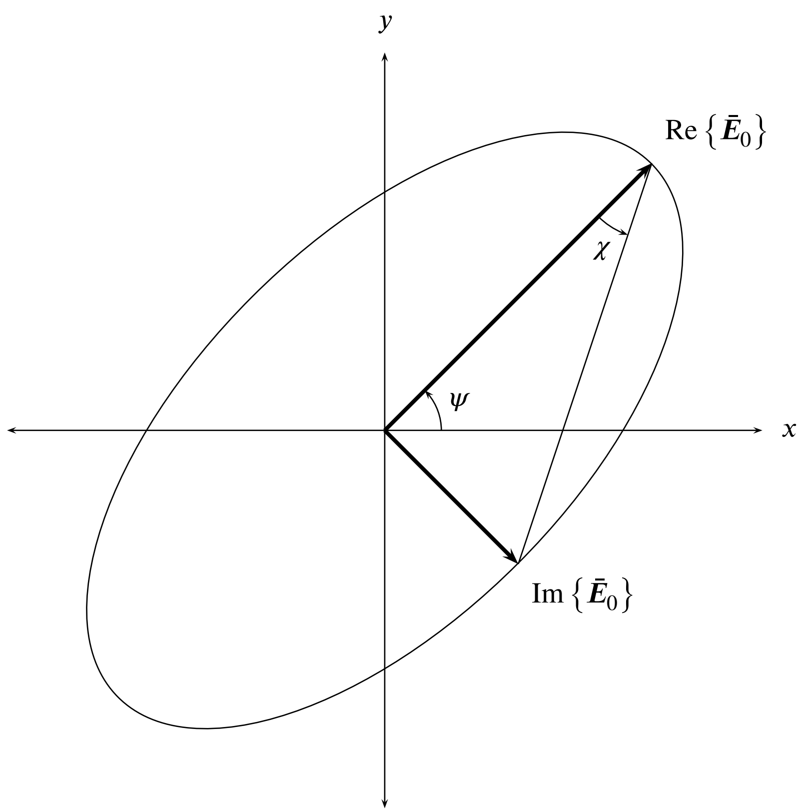

Figure 1: Geometry of the canonical polarization ellipse.

2.3. Generalization

Alas, equation 1 is glaringly limited—it models waves that necessarily propagate in free space along the $\hat{\boldsymbol{z}}$ axis, but in reality light may propagate along any given axis through any given medium. For the time being, we continue to assume that the wave is in free space, but we leave behind the assumption that the wave propagates along $\hat{\boldsymbol{z}}$. We suppose instead that the wave propagates along an arbitrary axis $\hat{\boldsymbol{l}}$ which is located in spherical coordinates by (θ, ϕ). Then, equation 1 generalizes to

$$\bar{\boldsymbol{E}}'(l,t) = A\mathrm{e}^{\mathrm{i}(kl-\omega t+\alpha)}\bar{\boldsymbol{E}}_0'(\theta,\phi,\chi,\psi) \tag{3}$$

where l is a distance along $\hat{\boldsymbol{l}}$ (in the same way that z is a distance along $\hat{\boldsymbol{z}}$) and $\bar{\boldsymbol{E}}_0'$ is the *generalized Jones vector*, or *GJV*, given by [1]. If we establish the transformation matrix

$$\boldsymbol{C}(\theta,\phi) = \begin{pmatrix} \sin\phi & \cos\theta\cos\phi & \sin\theta\cos\phi \\ -\cos\phi & \cos\theta\sin\phi & \sin\theta\sin\phi \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{4}$$

corresponding to the change of basis derived in [1] which maps $\hat{\boldsymbol{z}}$ to $\hat{\boldsymbol{l}}$, the generalized Jones vector may be written as

$$\bar{\boldsymbol{E}}_0'(\theta,\phi,\chi,\psi) = \boldsymbol{C}(\theta,\phi)\bar{\boldsymbol{E}}_0(\chi,\psi). \tag{5}$$

Simply put, a change of basis consistent with $\hat{\boldsymbol{l}}$ is applied to the canonical Jones vector to establish the generalized Jones vector. Moreover, the precise formulation of $\bar{\boldsymbol{E}}_0'$ which is exhibited by equation 7 of [1] may be obtained by performing the multiplication symbolically,

$$\begin{pmatrix} \cos\chi\,(\sin\phi\cos\psi + \cos\theta\cos\phi\sin\psi) + \mathrm{i}\sin\chi\,(\sin\phi\sin\psi - \cos\theta\cos\phi\cos\psi) \\ -\cos\chi\,(\cos\phi\cos\psi - \cos\theta\sin\phi\sin\psi) - \mathrm{i}\sin\chi\,(\cos\phi\sin\psi + \cos\theta\sin\phi\cos\psi) \\ -\sin\theta\,(\cos\chi\sin\psi - \mathrm{i}\sin\chi\cos\psi) \end{pmatrix}.$$

As mentioned in the previous section, equation 1 traces the canonical polarization ellipse in the plane perpendicular to $\hat{\boldsymbol{z}}$ (also known as the xy plane) when evaluated over z and t. Note that equation 3 traces the canonical polarization ellipse as well, though it does so in the plane perpendicular to $\hat{\boldsymbol{l}}$ when evaluated over l and t. In other words, we have reoriented the entire model so that it behaves with respect to $\hat{\boldsymbol{l}}$ as it did with respect to $\hat{\boldsymbol{z}}$.

2.4. Discussion

Radiance is the primary quantity of interest in the context of physically based rendering, and so it is sensible to determine the relationship between radiance and the generalized Jones vector. We recall that radiance, denoted $\mathrm{L}_{e,\Omega}$, measures power per unit solid angle per unit area,

$$\mathrm{L}_{e,\Omega} = \frac{\mathrm{W}}{\mathrm{sr}\cdot\mathrm{m}^2} = \frac{\mathrm{kg}}{\mathrm{sr}\cdot\mathrm{s}^3}$$

where the rightmost equality expresses the measure in base units (refer to tables 2 and 3 for a review of these units if necessary). Because the generalized Jones vector is decidedly monochromatic, the quantity we are truly interested in is *spectral* radiance, denoted $\mathrm{L}_{e,\Omega,\lambda}$, which measures radiance per unit wavelength,

$$\mathrm{L}_{e,\Omega,\lambda} = \frac{\mathrm{W}}{\mathrm{sr}\cdot\mathrm{m}^3} = \frac{\mathrm{kg}}{\mathrm{sr}\cdot\mathrm{m}\cdot\mathrm{s}^3}.$$

Then, upon integrating a quantity which measures spectral radiance over wavelength, we recover a quantity which measures radiance. As this suggests, we must sample over wavelength to implement the generalized Jones calculus, which is initially worrisome. Status quo rendering software presently utilizes some manner of "spectrum" class which stores multiple spectral radiance samples as floating point values to parallelize calculations—of course, this is no longer feasible when all calculations are wavelength-dependent and samples are complex vectors.

Before we address this concern, however, we must determine how measures of spectral radiance are extracted from the generalized Jones vector. To do this, let us introduce the *Poynting vector*. In free space as well as non-magnetic participating media, the Poynting vector $\boldsymbol{S}$ is given by

$$\boldsymbol{S} = \frac{\boldsymbol{E} \times \boldsymbol{B}}{\mu},$$

which measures the power per unit area of the electromagnetic field. Hence, it measures *irradiance*, which has units of $\mathrm{W} / \mathrm{m}^2$. We isolate a particular wavelength with the GJV by construction, so this quantity is in fact *spectral irradiance* in context with units $\mathrm{W} / \mathrm{m}^3$.

But the form of the Poynting vector shown above is the *instantaneous* Poynting vector. To obtain an appropriate measure of spectral irradiance as a scalar quantity, we may calculate the length of the *time-averaged* Poynting vector. For plane waves propagating in free space, which is likely a decent approximation for whichever medium the sensor is understood to be in, this is given by

$$\frac{1}{2\zeta} \| \bar{\boldsymbol{E}}' \|^2$$

where $\zeta = \mu c = 120\pi\,\Omega$ is the wave impedance of free space.[2] Upon scaling spectral irradiance by a measure of projected area (i.e., $\cos\theta$), we obtain spectral radiance.

Now, to address the concern of wavelength dependence. We contend that sampling over wavelength is indeed practicable. Figure 2 exhibits images produced by a wavelength dependent renderer whose path tracing algorithm is otherwise simplistic. While this renderer does not utilize the Jones calculus, it illustrates that the convergence rates of wavelength dependent algorithms are managable nevertheless—this is achieved by importance sampling wavelength from a weighted combination of the CIE XYZ color matching functions or, more specifically, the analytic fits of the CIE XYZ color matching functions given by Wyman and colleagues [6]. I have made the source available on GitHub, `github.com/imgsci/dense`.

[2] Note that Ω is the symbol for the SI unit of resistance, the ohm, and not a variable in this context. Refer to table 3.

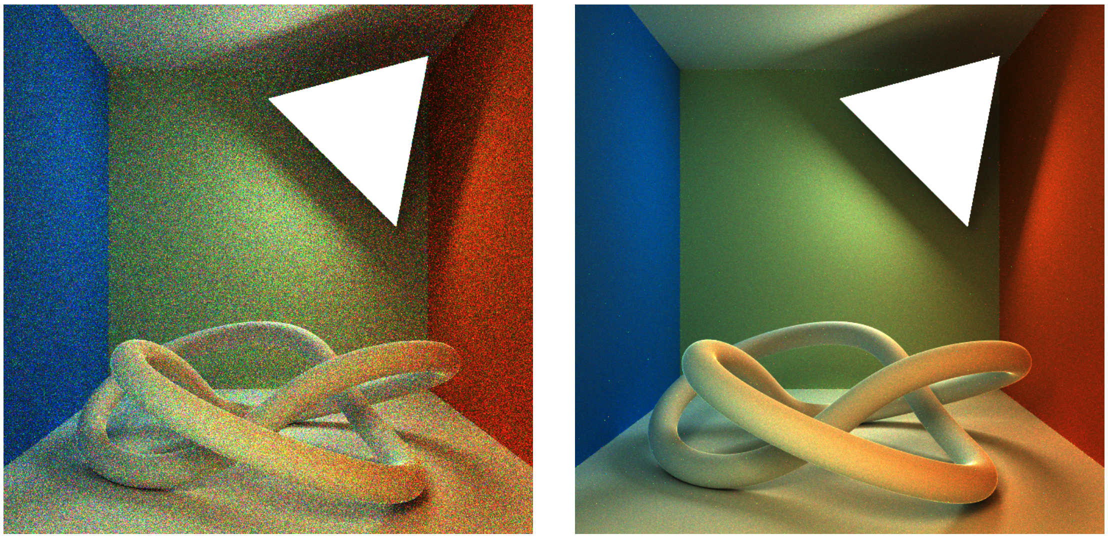

Figure 2: Wavelength dependent renders which have importance sampled wavelength from a weighted combination of the CIE XYZ color matching functions. On the left is the result of 16 samples per pixel. On the right is the result of 512 samples per pixel.

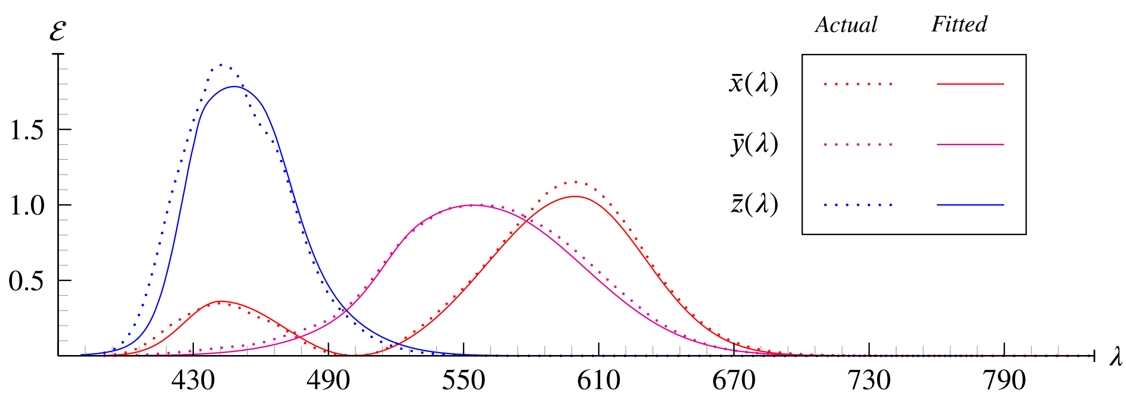

Figure 3: A graph plotting the XYZ color matching functions (which are tabulated and thus have no closed form) against the analytic fits of Wyman and collegues [6]. It is easy to see that the approximations are quite good! Above, λ denotes wavelength measured in nanometers and $\mathcal{E}$ indicates the response of the human eye.

3. The Generalized Jones Matrix

It is likewise to be observed, that this society has a peculiar cant and jargon of their own, that no other mortal can understand, and wherein all their laws are written, which they take special care to multiply;

—Jonathan Swift, *Gulliver's Travels*

3.1. Context and Functionality

We move on to the *generalized Jones matrix*, or *GJM*, whose theoretical construction is fleshed out in [2]. Establishing the canonical Jones vector is a natural step in developing its generalization, but the same cannot be said about the Jones matrix. Make no mistake, there is indeed a canonical Jones matrix which deals exclusively with canonical Jones vectors, but we pay it no mind—equipped with the GJV, the GJM is sensibly developed from scratch.

Per equations 1 and 3, we have assumed thus far that the electric field of a plane wave is a function of two scalars, the first of which represents a distance along the propagation axis and the second of which represents a point in time. But the definition of a plane wave suggests that it should be defined at all points in space, not just those points that are colinear with the propagation axis. We have also assumed that the wave propagates though free space.

In the interest of foregoing these assumptions, let $\boldsymbol{r}$ denote an arbitrary spatial coordinate and let $\boldsymbol{k} = k\hat{\boldsymbol{l}}$ denote the wavevector (the propagation axis scaled by the angular wavenumber). Notice that $\boldsymbol{k} \cdot \boldsymbol{r}$ is the projected distance along $\hat{\boldsymbol{l}}$ of each $\boldsymbol{r}$, multiplied by k. Then, equation 3 generalizes to

$$\bar{\boldsymbol{E}}''(\boldsymbol{r}, t) = \mathrm{e}^{-\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}\boldsymbol{\Gamma}(\boldsymbol{k}\cdot\boldsymbol{r})}\bar{\boldsymbol{E}}_0'' \tag{6}$$

where $\bar{\boldsymbol{E}}_0'' = A\mathrm{e}^{\mathrm{i}\alpha}\,\bar{\boldsymbol{E}}_0'(\theta, \phi, \chi, \psi)$ is the initial state of the electric field and $\boldsymbol{\Gamma}$ denotes the *differential generalized Jones matrix*, or *dGJM*, which characterizes a necessarily homogeneous but potentially anisotropic participating medium. As established in [2], the dGJM is a complex linear combination of the Gell-Mann matrices

$$\boldsymbol{M}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \boldsymbol{M}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{M}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{M}_3 = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{M}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \cdot \frac{1}{\sqrt{3}} \qquad \boldsymbol{M}_5 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{M}_6 = \begin{pmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{pmatrix} \qquad \boldsymbol{M}_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \boldsymbol{M}_8 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}$$

which are the three dimensional analogues of the Pauli spin matrices. However, the definition of the dGJM given in this paper,

$$\boldsymbol{\Gamma} = \frac{1}{2}\sum_{i=0}^{8} m_i \boldsymbol{M}_i, \tag{7}$$

is not identical to the definition given in [2], which includes a factor of i. We state this factor of i in equation 6 to remain consistent with equations 1 and 3. In any case, the coefficients are

$$\begin{pmatrix} m_0 & m_1 & \cdots & m_8 \end{pmatrix} = \begin{pmatrix} 2\gamma & \gamma_q^{xy} & \gamma_u^{xy} & \gamma_v^{xy} & 2\gamma_q^z/\sqrt{3} & \gamma_u^{xz} & \gamma_v^{xz} & \gamma_u^{yz} & \gamma_v^{yz} \end{pmatrix}$$

where $\gamma = \eta + \mathrm{i}\kappa$ is the isotropic propagation constant, with refractive index η and extinction κ, and the similarly defined $\gamma_{q,u,v} = \eta_{q,u,v} + \mathrm{i}\kappa_{q,u,v}$ are anisotropic propagation constants which characterize the participating medium. As noted in [2], γ_q^{xy} quantifies the linear birefringence and dichroism of the xy plane, γ_q^z quantifies the difference in linear retardance and absorption between the z direction and the xy plane, and the $\gamma_{u,v}$ quantify the linear $\pm 45°$ and circular birefringence and dichroism of the xy, xz, and yz planes.

The generalized Jones matrix itself is given by $\mathrm{e}^{\mathrm{i}\boldsymbol{\Gamma}(\boldsymbol{k}\cdot\boldsymbol{r})}$. Since the generalized Jones matrices are related to the differential generalized Jones matrices by an exponential map, the generalized Jones matrices form a *Lie group* whose associated *Lie algebra* consists of the differential generalized Jones matrices. Hence, we may think of dGJMs as infinitesimal generators of GJMs. Although we refrain from delving further into the Lie theoretical approach to the Jones matrix, we acknowledge that this approach offers much insight yet unexhibited. See [3] and [5] for a brief introduction to Lie theory.

3.2. Microscopic Field Equations

Differentiating equation 6 with respect to space and time is straightforward (after writing out the dot product in the exponential). Taking derivatives repeatedly, we find that

$$\frac{\partial^n}{\partial r_i^n}\bar{\boldsymbol{E}}'' = \left(\mathrm{i}k_i\boldsymbol{\Gamma}\right)^n \bar{\boldsymbol{E}}'' \text{ and } \frac{\partial^n}{\partial t^n}\bar{\boldsymbol{E}}'' = (-\mathrm{i}\omega)^n\, \bar{\boldsymbol{E}}''.$$

Plugging in the coordinate derivatives, we discover closed form expressions for the usual differential operators on the electric field,

$$\begin{aligned} \nabla \cdot \bar{\boldsymbol{E}}'' &= \boldsymbol{k} \cdot \mathrm{i}\boldsymbol{\Gamma}\bar{\boldsymbol{E}}'', \\ \nabla \times \bar{\boldsymbol{E}}'' &= \boldsymbol{k} \times \mathrm{i}\boldsymbol{\Gamma}\bar{\boldsymbol{E}}'', \\ \nabla^2 \bar{\boldsymbol{E}}'' &= -k^2\boldsymbol{\Gamma}^2\bar{\boldsymbol{E}}'', \text{ and} \\ \Box^2 \bar{\boldsymbol{E}}'' &= -k^2\boldsymbol{\Gamma}^2\bar{\boldsymbol{E}}'' + k^2\bar{\boldsymbol{E}}''. \end{aligned}$$

To satisfy Maxwell's equations, the corresponding magnetic field vector must be given by

$$\bar{\boldsymbol{B}}''(\boldsymbol{r},t) = \frac{\boldsymbol{k} \times \boldsymbol{\Gamma}\bar{\boldsymbol{E}}''(\boldsymbol{r},t)}{\omega}$$

which is reminiscent of the magnetic field vector corresponding to equation 1 established in section 2.2. Taking yet more coordinate derivatives and applying vector identities, we discover closed form expressions for the usual differential operators on the magnetic field in terms of the electric field,

$$\begin{aligned} \nabla \cdot \bar{\boldsymbol{B}}'' &= 0, \\ \nabla \times \bar{\boldsymbol{B}}'' &= \frac{1}{\omega}\left(\boldsymbol{k} \times \left(\boldsymbol{k} \times \mathrm{i}\boldsymbol{\Gamma}^2\bar{\boldsymbol{E}}''\right)\right), \\ \nabla^2 \bar{\boldsymbol{B}}'' &= \frac{1}{\omega}\left(\boldsymbol{k} \times \left(\boldsymbol{k} \times \left(\boldsymbol{k} \times \boldsymbol{\Gamma}^3\bar{\boldsymbol{E}}''\right)\right)\right), \text{ and} \\ \Box^2 \bar{\boldsymbol{B}}'' &= \frac{1}{\omega}\left(\boldsymbol{k} \times \left(\boldsymbol{k} \times \left(\boldsymbol{k} \times \boldsymbol{\Gamma}^3\bar{\boldsymbol{E}}''\right)\right) - k^2\left(\boldsymbol{k} \times \boldsymbol{\Gamma}\bar{\boldsymbol{E}}''\right)\right). \end{aligned}$$

Expressions for the induced charge density $\bar{\rho}''$ and current density vector $\bar{\boldsymbol{J}}''$ follow suit,

$$\begin{aligned}\bar{\rho}'' &= \left(\boldsymbol{k}\cdot \mathrm{i}\boldsymbol{\Gamma}\bar{\boldsymbol{E}}''\right)\varepsilon \text{ and}\\ \bar{\boldsymbol{J}}'' &= \left(\boldsymbol{k}\times\left(\boldsymbol{k}\times\boldsymbol{\Gamma}^2\bar{\boldsymbol{E}}''\right)+k^2\bar{\boldsymbol{E}}''\right)\frac{\mathrm{i}}{\omega\mu},\end{aligned}$$

which vanish in free space (where $\boldsymbol{\Gamma}$ is the identity matrix and the electric field $\bar{\boldsymbol{E}}''$ is perpendicular to the wavevector $\boldsymbol{k}$) as we expect.

3.3. Computing the GJM

We must be able to exponentiate arbitrary dGJMs to calculate arbitrary GJMs. In particular, we must evaluate expressions of the form $\mathrm{e}^{\mathrm{i}l\boldsymbol{\Gamma}}$, where l is a real number and $\boldsymbol{\Gamma}$ is defined as above in equation 7. To that end, let us establish the *differential anisotropy matrix*

$$\boldsymbol{A}=\frac{1}{2}\sum_{k=1}^{8}m_k\boldsymbol{M}_k=\frac{1}{2}\begin{pmatrix}\gamma_q^{xy}+\frac{2}{3}\gamma_q^{z} & \gamma_u^{xy}-\mathrm{i}\gamma_v^{xy} & \gamma_u^{xz}-\mathrm{i}\gamma_v^{xz}\\ \gamma_u^{xy}+\mathrm{i}\gamma_v^{xy} & -\gamma_q^{xy}+\frac{2}{3}\gamma_q^{z} & \gamma_u^{yz}-\mathrm{i}\gamma_v^{yz}\\ \gamma_u^{xz}+\mathrm{i}\gamma_v^{xz} & \gamma_u^{yz}+\mathrm{i}\gamma_v^{yz} & -\frac{4}{3}\gamma_q^{z}\end{pmatrix} \tag{8}$$

so that $\boldsymbol{\Gamma}$ may be rewritten $\gamma\boldsymbol{I}+\boldsymbol{A}$ where $\gamma\boldsymbol{I}$ and $\boldsymbol{A}$ may be interpreted as the isotropic and anisotropic terms respectively. Next, let us rewrite the exponential as

$$\mathrm{e}^{\mathrm{i}l\boldsymbol{\Gamma}}=\mathrm{e}^{\mathrm{i}l(\gamma\boldsymbol{I}+\boldsymbol{A})}=\mathrm{e}^{\mathrm{i}l\gamma\boldsymbol{I}}\mathrm{e}^{\mathrm{i}l\boldsymbol{A}}=\mathrm{e}^{\mathrm{i}l\gamma}\boldsymbol{I}\mathrm{e}^{\mathrm{i}l\boldsymbol{A}}=\mathrm{e}^{\mathrm{i}l\gamma}\mathrm{e}^{\mathrm{i}l\boldsymbol{A}}.$$

Note that this is only “legal” because $\gamma\boldsymbol{I}$ and $\boldsymbol{A}$ commute—in the general case, the matrix exponential of a sum is not equal to the product of the exponentials of the terms. In this case, however, $\gamma\boldsymbol{I}$ is a scalar multiple of the identity matrix which commutes universally, so the equality holds.

Calculating $\mathrm{e}^{\mathrm{i}l\gamma}$ is straightforward (is accomplished by a standard library call), so the challenge lies only in calculating $\mathrm{e}^{\mathrm{i}l\boldsymbol{A}}$. By the Cayley-Hamilton theorem, we know that there exist scalars which express $\mathrm{e}^{\mathrm{i}l\boldsymbol{A}}$ as a linear combination of the first three powers of $\boldsymbol{A}$. That is,

$$\mathrm{e}^{\mathrm{i}l\boldsymbol{A}}=\alpha_0\boldsymbol{I}+\alpha_1\boldsymbol{A}+\alpha_2\boldsymbol{A}^2$$

where the α_j are determined by a linear system in the exponentials of the eigenvalues λ_j of $\boldsymbol{A}$. When the eigenvalues of $\boldsymbol{A}$ are distinct such that $\lambda_1\neq\lambda_2\neq\lambda_3$, the straightforward system

$$\begin{pmatrix}1 & \lambda_1 & (\lambda_1)^2\\ 1 & \lambda_2 & (\lambda_2)^2\\ 1 & \lambda_3 & (\lambda_3)^2\end{pmatrix}\begin{pmatrix}\alpha_0\\ \alpha_1\\ \alpha_2\end{pmatrix}=\begin{pmatrix}\mathrm{e}^{\mathrm{i}l\lambda_1}\\ \mathrm{e}^{\mathrm{i}l\lambda_2}\\ \mathrm{e}^{\mathrm{i}l\lambda_3}\end{pmatrix} \tag{9}$$

is sufficient to determine the α_j. However, the system in equation 9 is singular when any eigenvalue is repeated. If λ_i is unique (has multiplicity 1) and λ_j is repeated (has multiplicity 2), then we may solve a modified system,

$$\begin{pmatrix}1 & \lambda_i & (\lambda_i)^2\\ 1 & \lambda_j & (\lambda_j)^2\\ 0 & 1 & 2\lambda_j\end{pmatrix}\begin{pmatrix}\alpha_0\\ \alpha_1\\ \alpha_2\end{pmatrix}=\begin{pmatrix}\mathrm{e}^{\mathrm{i}l\lambda_i}\\ \mathrm{e}^{\mathrm{i}l\lambda_j}\\ \mathrm{i}l\mathrm{e}^{\mathrm{i}l\lambda_j}\end{pmatrix}, \tag{10}$$

where we have taken the derivative of the last equation with respect to λ_j to guarantee that the system is non-singular. In the event that $\lambda_1 = \lambda_2 = \lambda_3$, we may perform a similar process wherein we take the first derivative of the second equation and the second derivative of the third equation.

First recognize that $\boldsymbol{A}$ is *traceless* in that its trace (the sum of its diagonal entries) is zero by construction. Given that the trace of every matrix is the sum of its eigenvalues, if the eigenvalues of a traceless matrix are equivalent, they are equivalent to zero. Therefore, $\lambda_1 = \lambda_2 = \lambda_3$ implies $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Taking the first derivative of the second equation and the second derivative of the third equation, then substituting zero for each eigenvalue, we arrive at

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{11}$$

which is easily solved in closed form. But before we solve any systems, we must find the eigenvalues of course. Again leveraging the observation that $\boldsymbol{A}$ is traceless, we know its characteristic polynomial has a vastly simplified form,

$$P_{\boldsymbol{A}}(\lambda) = \lambda^3 - \lambda \frac{\operatorname{tr} \boldsymbol{A}^2}{2} - \frac{\operatorname{tr} \boldsymbol{A}^3}{3},$$

whose roots are not particularly difficult to compute. We simply let

$$u = \sqrt[3]{r + \sqrt{r^2 - q^3}}$$

where

$$q = \frac{\operatorname{tr} \boldsymbol{A}^2}{6} \text{ and } r = \frac{\operatorname{tr} \boldsymbol{A}^3}{6}.$$

If u is zero, the eigenvalues are all zero. Otherwise, let $v = q \,/\, u$, and the eigenvalues are

$$\begin{aligned} \lambda_1 &= u + v, \\ \lambda_2 &= -\frac{(1 + \mathrm{i}\sqrt{3})u + (1 - \mathrm{i}\sqrt{3})v}{2}, \text{ and} \\ \lambda_3 &= -\frac{(1 - \mathrm{i}\sqrt{3})u + (1 + \mathrm{i}\sqrt{3})v}{2}. \end{aligned}$$

Implementation

An implementation of this algorithm in C++14 is provided below. Although it is not stand-alone, the full (header-only) source is available on GitHub, `github.com/imgsci/foyer`. Definitions of the data types in use are not provided, but they are meaningfully named and their precise implementation is not crucial to the success of the algorithm. Nonetheless, it is worth noting that the `FieldT` type is a complex floating point type, the `FloatT` type is the underlying floating point type, the `VectorT` type is a three dimensional complex vector, and the `MatrixT` type is a three-by-three dimensional complex matrix.

```
// Differential anisotropy matrix.
MatrixT Alpha() const
{
  return {
    gammaq.dz / FloatT(3.0) + gammaq.xy * FloatT(0.5),
    gammau.xy * FloatT(0.5) - gammav.xy * FloatT(0.5) * FieldT(0.0, 1.0),
    gammau.xz * FloatT(0.5) - gammav.xz * FloatT(0.5) * FieldT(0.0, 1.0),
    gammau.xy * FloatT(0.5) + gammav.xy * FloatT(0.5) * FieldT(0.0, 1.0),
    gammaq.dz / FloatT(3.0) - gammaq.xy * FloatT(0.5),
    gammau.yz * FloatT(0.5) - gammav.yz * FloatT(0.5) * FieldT(0.0, 1.0),
    gammau.xz * FloatT(0.5) + gammav.xz * FloatT(0.5) * FieldT(0.0, 1.0),
    gammau.yz * FloatT(0.5) + gammav.yz * FloatT(0.5) * FieldT(0.0, 1.0),
    gammaq.dz / FloatT(3.0) * FloatT(-2.0)
  };
}

// Exponential of differential Jones matrix.
MatrixT expiGamma(FloatT l) const
{
  constexpr FloatT eps = epsilon<FloatT>;
  constexpr FieldT const_1 = FieldT(0.5,  0.86602540378443864676);
  constexpr FieldT const_2 = FieldT(0.5, -0.86602540378443864676);

  // Let phi = i * l.
  FieldT phi = FieldT(0, l);

  // Compute powers of Alpha.
  MatrixT Alpha0 = MatrixT(1);
  MatrixT Alpha1 = Alpha();
  MatrixT Alpha2 = Alpha1 * Alpha1;
  MatrixT Alpha3 = Alpha2 * Alpha1;

  // Declare solutions to the exponential eigenvalue problem.
  VectorT x;

  // Declare eigenvalues.
  VectorT lambda;

  // Intermediate steps to roots.
  FieldT q = Alpha2.trace() / FloatT(6.0);
  FieldT r = Alpha3.trace() / FloatT(6.0);
  FieldT u = cbrt(r + sqrt(r * r - q * q * q));
  if ((eps * abs(q)) < abs(u)) {

    // Solve for the roots, being the eigenvalues of Alpha.
    FieldT v = q / u;
    lambda[0] = u + v;
    lambda[1] = -(u * const_1 + v * const_2);
    lambda[2] = -(u * const_2 + v * const_1);
```

```
    // Initialize linear system for exponential.
    MatrixT M;
    VectorT b;
    if (abs(lambda[1] - lambda[2]) < FloatT(1e-5)) {
      // There is a repeated eigenvalue, set up the modified system.
      M[0] = {FloatT(1), lambda[0], lambda[0] * lambda[0]};
      M[1] = {FloatT(1), lambda[1], lambda[1] * lambda[1]};
      M[2] = {FloatT(0), FloatT(1), FloatT(2) * lambda[1]};
      b[0] = exp(phi * lambda[0]);
      b[1] = exp(phi * lambda[1]);
      b[2] = phi * b[1];
    }
    else if (abs(lambda[0] - lambda[1]) < FloatT(1e-5)) {
      // There is a repeated eigenvalue, set up the modified system.
      M[0] = {FloatT(1), lambda[2], lambda[2] * lambda[2]};
      M[1] = {FloatT(1), lambda[0], lambda[0] * lambda[0]};
      M[2] = {FloatT(0), FloatT(1), FloatT(2) * lambda[0]};
      b[0] = exp(phi * lambda[2]);
      b[1] = exp(phi * lambda[0]);
      b[2] = phi * b[1];
    }
    else if (abs(lambda[0] - lambda[2]) < FloatT(1e-5)) {
      // There is a repeated eigenvalue, set up the modified system.
      M[0] = {FloatT(1), lambda[1], lambda[1] * lambda[1]};
      M[1] = {FloatT(1), lambda[0], lambda[0] * lambda[0]};
      M[2] = {FloatT(0), FloatT(1), FloatT(2) * lambda[0]};
      b[0] = exp(phi * lambda[1]);
      b[1] = exp(phi * lambda[0]);
      b[2] = phi * b[1];
    }
    else {
      // All eigenvalues are distinct, set up the usual system.
      M[0] = {FieldT(1), lambda[0], lambda[0] * lambda[0]};
      M[1] = {FieldT(1), lambda[1], lambda[1] * lambda[1]};
      M[2] = {FieldT(1), lambda[2], lambda[2] * lambda[2]};
      b[0] = exp(phi * lambda[0]);
      b[1] = exp(phi * lambda[1]);
      b[2] = exp(phi * lambda[2]);
    }

    // Done
    x = inverse(M) * b;
  }
  else {
    // All eigenvalues are zero, solve the system directly.
    x[0] = FloatT(1.0);
    x[1] = FloatT(1.0) * phi;
    x[2] = FloatT(0.5) * phi * phi;
  }

  // Isotropic exponential times anisotropic exponential.
  return exp(phi * gamma) * (Alpha0 * x[0] + Alpha1 * x[1] + Alpha2 * x[2]);
}
```

3.4. Discussion

There are a few auxillary remarks worth making. Firstly, the structure of the differential anisotropy matrix given by equation 8 is rather elegant. To see why, recall that the real and imaginary parts of the isotropic propagation constant γ are the linear retardance η and the absorption κ—the anisotropy matrix $\boldsymbol{A}$ has a similar factorization. It is well known that Hermitian matrices are analogous to real numbers while anti-Hermitian matrices are analogous to imaginary numbers. Decomposing $\boldsymbol{A}$ into its Hermitian and anti-Hermitian parts,

$$\boldsymbol{H} = \frac{\boldsymbol{A} + \boldsymbol{A}^\dagger}{2} \text{ and}$$
$$\boldsymbol{K} = \frac{\boldsymbol{A} - \boldsymbol{A}^\dagger}{2},$$

we notice a striking resemblence to the isotropic case—the Hermitian part $\boldsymbol{H}$ and the anti-Hermitian part $\boldsymbol{K}$ correspond to the real and imaginary parts of the anisotropic propagation constants $\gamma_{q,u,v}$ such that $\boldsymbol{H}$ gathers the birefringence coefficients,

$$\boldsymbol{H} = \frac{1}{2}\begin{pmatrix} \eta_q^{xy} + \frac{2}{3}\eta_q^z & \eta_u^{xy} - \mathrm{i}\eta_v^{xy} & \eta_u^{xz} - \mathrm{i}\eta_v^{xz} \\ \eta_u^{xy} + \mathrm{i}\eta_v^{xy} & -\eta_q^{xy} + \frac{2}{3}\eta_q^z & \eta_u^{yz} - \mathrm{i}\eta_v^{yz} \\ \eta_u^{xz} + \mathrm{i}\eta_v^{xz} & \eta_u^{yz} + \mathrm{i}\eta_v^{yz} & -\frac{4}{3}\eta_q^z \end{pmatrix},$$

and $\boldsymbol{K}$ gathers the dichroism coefficients,

$$\boldsymbol{K} = \frac{\mathrm{i}}{2}\begin{pmatrix} \kappa_q^{xy} + \frac{2}{3}\kappa_q^z & \kappa_u^{xy} - \mathrm{i}\kappa_v^{xy} & \kappa_u^{xz} - \mathrm{i}\kappa_v^{xz} \\ \kappa_u^{xy} + \mathrm{i}\kappa_v^{xy} & -\kappa_q^{xy} + \frac{2}{3}\kappa_q^z & \kappa_u^{yz} - \mathrm{i}\kappa_v^{yz} \\ \kappa_u^{xz} + \mathrm{i}\kappa_v^{xz} & \kappa_u^{yz} + \mathrm{i}\kappa_v^{yz} & -\frac{4}{3}\kappa_q^z \end{pmatrix}.$$

It is also well known that linear combinations of the Gell-Mann matrices with real coefficients form a basis for the infinitesimal generators of the special unitary group. Referring to above factorization, it is no surprise that $\mathrm{i}l\boldsymbol{H}$, which is traceless anti-Hermitian and is thus guaranteed to have a special unitary exponential, remains when all absorption coefficients vanish.

Alas, this factorization does not help with computing the exponential in the general case—that is, we cannot split the exponential further into $\mathrm{e}^{\mathrm{i}l\boldsymbol{H}}\mathrm{e}^{\mathrm{i}l\boldsymbol{K}}$ because $\boldsymbol{H}$ and $\boldsymbol{K}$ do not necessarily commute. In particular,

$$[\boldsymbol{H}, \boldsymbol{K}] = \frac{1}{2}[\boldsymbol{A}^\dagger, \boldsymbol{A}],$$

which implies that $\boldsymbol{H}$ and $\boldsymbol{K}$ commute if and only if $\boldsymbol{A}$ is normal.

4. Future Prospects

In solving a problem of this sort, the grand thing is to be able to reason backward.

—Sir Arthur Conan Doyle, *A Study In Scarlet*

In summary, the generalized Jones calculus offers an elegant model for potentially inhomogeneous monochromatic plane waves which propagate through potentially anisotropic media along arbitrary axes—this concerns a tremendous class of problems in optics, physics, and crystallography. To take full advantage of this modeling technique in the future, the theoretical footing of the generalized Jones calculus must be strengthened. Moreover, we must uncover precisely how generalized Jones vectors and matrices relate to pre-existing theory. For instance, the differential generalized Jones matrix $\boldsymbol{\Gamma}$ as it appears in equation 6 functions as a propagation constant, which suggests by analogy to the scalar case that its square should be equal to the product of the permeability and permittivity tensors of the medium in question. To clarify, let us briefly recall the isotropic case.

For an isotropic medium $\mathcal{M}$, the square of its propagation constant $\gamma_{\mathcal{M}}$ is equal to the product of its relative permeability $\mu_{\mathcal{M}}$ and its relative permittivity $\varepsilon_{\mathcal{M}}$. So,

$$\gamma_{\mathcal{M}}^2 = \mu_{\mathcal{M}}\varepsilon_{\mathcal{M}}.$$

However, the relative permeability $\boldsymbol{\mu}_{\mathcal{N}}$ and relative permittivity $\boldsymbol{\varepsilon}_{\mathcal{N}}$ of an anisotropic medium $\mathcal{N}$ are tensors (specifically, three-by-three dimensional complex Hermitian matrices). Knowing that the differential generalized Jones matrix $\boldsymbol{\Gamma}_{\mathcal{N}}$ which characterizes the same medium consists of supposed propagation constants and functions as a propagation constant itself, it is reasonable to suspect

$$\boldsymbol{\Gamma}_{\mathcal{N}}^2 = \boldsymbol{\mu}_{\mathcal{N}}\boldsymbol{\varepsilon}_{\mathcal{N}}.$$

This is conjecture at the moment, as I have not managed to determine a sensible relationship between the coefficients appearing in equation 7 and the unique entries of $\boldsymbol{\mu}_{\mathcal{N}}$ and $\boldsymbol{\varepsilon}_{\mathcal{N}}$. Even so, it follows that the macroscopic formulations of the microscopic field equations in section 3.2 for plane waves in anisotropic media are consistent with the established macroscopic field equations for plane waves in isotropic media if we suppose outright that such a relationship exists.

A plethora of open problems remain unaddressed even strictly in the context of graphics. Among topics elided from the discussion here are the generalizations of Snell's law and Fresnel's equations, which are non-trivial. To derive the reflected and transmitted waves spawned by equation 6 at the boundary between potentially anisotropic media, the *interface conditions* for the electromagnetic field must be taken into account. Further research should also examine the efficacy of more intricate rendering algorithms and variance reduction techniques, such as bidirectional Markov Chain Monte Carlo methods, for wavelength dependent rendering. In the same vein, I wonder as to whether there is an optimal weighting of the CIE XYZ color matching functions for importance sampling. It is better to have too much to do than not enough, I suppose.

References

[1] R. M. A. Azzam. “Three-dimensional polarization states of monochromatic light fields”. In: *J. Opt. Soc. Am. A* 28.11 (Nov. 2011), pp. 2279–2283.

[2] Noé Ortega-Quijano and José Luis Arce-Diego. “Generalized Jones matrices for anisotropic media”. In: *Opt. Express* 21.6 (Mar. 2013), pp. 6895–6900.

[3] Todd Rowland. *Lie Group*. url: `http://mathworld.wolfram.com/LieGroup.html`.

[4] Ambler Thompson and Barry N. Taylor. *Guide for the Use of the International System of Units (SI)*. Gaithersburg, MD 20899: National Institute of Standards and Technology, Mar. 2008.

[5] Eric W. Weisstein. *Lie Algebra*. url: `http://mathworld.wolfram.com/LieAlgebra.html`.

[6] Chris Wyman, Peter-Pike Sloan, and Peter Shirley. “Simple Analytic Approximations to the CIE XYZ Color Matching Functions”. In: *Journal of Computer Graphics Techniques (JCGT)* 2.2 (2013), pp. 1–11.