STOKES PARAMETERS AND GIBBS BIVECTORS

Michael Hayes

Abstract For a single monochromatic wave train, it is shown that the use of Gibbs bivectors leads in a natural and economical way to the introduction of Stokes parameters.

Introduction

Stokes, in a study of light waves published [1] in 1852, introduced what are now called 'Stokes parameters.' These are functions only of the electromagnetic wave. The polarization state of a beam of light (either natural, totally or partially polarized) can be described in terms of these four parameters [2, 3]. The purpose of this note is to show how the use of Gibbs bivectors [4, 5] leads to a direct and economical way of introducing Stokes parameters.

First, a simple observation. Any vector $\mathbf{a}$ (say) lying in a plane may be represented in an infinity of ways as a linear combination of two arbitrary vectors, $\mathbf{b}$ and $\mathbf{c}$ (say), in the plane, provided of course that $\mathbf{b}$ and $\mathbf{c}$ are not parallel. Thus $\mathbf{a} = \beta\mathbf{b} + \gamma\mathbf{c}$, for some scalars β and γ. Provided the vectors $\mathbf{b}$ and $\mathbf{c}$ are chosen to be orthogonal, then the squared length of $\mathbf{a}$, namely $a^2 = \mathbf{a}\cdot\mathbf{a}$, may be written simply as $a^2 = \beta^2 b^2 + \gamma^2 c^2$.

It is the generalization of this simple observation to the case of complex vectors, or bivectors, which leads naturally to Stokes parameters.

Here, Gibbs bivectors are introduced and some of their properties presented. Then a single monochromatic train of elliptically polarized plane transverse waves is considered.

Background

In 1853, the year following the publication of Stokes' paper [1], Hamilton, [6], in the context of quaternions, coined the word "bivector" for the combination $\mathbf{a}+i\mathbf{b}$, where $\mathbf{a}$ and $\mathbf{b}$ are real vectors. Complex numbers $\alpha+i\beta$ had been called biscalars! Hamilton seems not to have ever used bivectors. The theory was developed by Gibbs, [4], in 1881, 1884. He presented seven pages on bivectors in his seventy page pamphlet "Elements of Vector Analysis," which laid the modern foundations of vector algebra. Gibbs printed and circulated this privately. The work on bivectors was generally ignored, possibly because in parts it is difficult to read. The phrase "too condensed and too difficult" which was used by Lord Rayleigh, [7, page xiv], in writing to Gibbs about his famous paper "Equilibrium of Heterogeneous Substances" is apposite here also.

Gibbs recognized that an ellipse could be associated with a bivector. He naturally used bivectors in the description of elliptically polarized electromagnetic waves. It is here that the connection with Stokes parameters is made.

Bivectors

A pair of orthogonal radii to a circle are said to be 'conjugate.' If the circle is drawn on a sheet of rubber which is then stretched uniformly, the circle becomes an ellipse and pairs of conjugate radii of the circle become conjugate radii to the ellipse. The property that the tangent to the circle at the tip of the radius is parallel to the conjugate carries over to the ellipse. If $\mathbf{i}$ and $\mathbf{j}$ are parallel unit vectors then the position vector $\mathbf{r}$ given by

$$\mathbf{r} = a(\mathbf{i}\cos\theta + \mathbf{j}\sin\theta)$$

describes a circle of radius a. Also, if $\mathbf{a}$ and $\mathbf{b}$ is any pair of non-parallel vectors, then

$$\mathbf{r} = \mathbf{a}\cos\theta + \mathbf{b}\sin\theta$$

describes an ellipse in which $\mathbf{a}$ and $\mathbf{b}$ are conjugate radii. The tangent $d\mathbf{r}/d\theta$ at $\theta = 0$ $(\pi/2)$ is parallel to $\mathbf{a}$ $(\mathbf{b})$. The pair $\mathbf{r}(\theta)$ and $\mathbf{r}(\theta + \pi/2)$ are conjugate, [5].

The combination $\mathbf{D} = \mathbf{a} + i\mathbf{b}$ is said to be a bivector. Throughout capital bold face letters $\mathbf{D}$, $\mathbf{E}$, ..., are used to denote bivectors. The real and imaginary parts of the bivector $\mathbf{D}$ are denoted by $(\mathbf{D})^+$ and $(\mathbf{D})^-$, respectively. Associated with $\mathbf{D}$ is a unique ellipse, $\mathbf{r} = \mathbf{a}\cos\theta + \mathbf{b}\sin\theta$, $(0 \leq \theta \leq 2\pi)$ and a sense of description, from $(\mathbf{D})^+ = \mathbf{a}$ to $(\mathbf{D})^- = \mathbf{b}$. The complex conjugate of $\mathbf{D}$ is $\bar{\mathbf{D}} = \mathbf{a} - i\mathbf{b}$. The bivectors $\mathbf{D}$ and $\bar{\mathbf{D}}$ have the same ellipse associated with them, but in the case of $\bar{\mathbf{D}}$ the sense of description is from $(\bar{\mathbf{D}})^+ = \mathbf{a}$ to $(\bar{\mathbf{D}})^- = -\mathbf{b}$, which is opposite to that of the ellipse of $\mathbf{D}$.

Two bivectors $\mathbf{D}$ and $\mathbf{E}$ are said to be **parallel** if there exists a scalar, λ, such that $\mathbf{D} = \lambda\mathbf{E}$. Otherwise they are linearly independent.

The dot product of the bivectors $\mathbf{D} = \mathbf{a} + i\mathbf{b}$ and $\mathbf{E} = \mathbf{p} + i\mathbf{q}$ is defined in the usual way:

$$\mathbf{D} \cdot \mathbf{E} = \mathbf{a} \cdot \mathbf{p} - \mathbf{b} \cdot \mathbf{q} + i(\mathbf{b} \cdot \mathbf{p} + \mathbf{a} \cdot \mathbf{q}).$$

If $\mathbf{D} \cdot \mathbf{E} = 0$, then $\mathbf{D}$ and $\mathbf{E}$ are said to be **orthogonal**.

If $\mathbf{D} \cdot \mathbf{D} = 0$, then $\mathbf{a} \cdot \mathbf{a} = \mathbf{b} \cdot \mathbf{b}$, $\mathbf{a} \cdot \mathbf{b} = 0$ so that in this case the ellipse of $\mathbf{D}$ is a circle. For example, if $\mathbf{D} = \mathbf{i} + i\mathbf{j}$, then $\mathbf{D} \cdot \mathbf{D} = 0$.

The "intensity" of $\mathbf{D}$ is defined to be $\mathbf{D} \cdot \bar{\mathbf{D}}$. A bivector $\mathbf{D}$ of unit intensity may be represented by

$$\mathbf{D} = \cos\beta\,\mathbf{r} + i\sin\beta\,\mathbf{s}, \tag{1}$$

where $\mathbf{r}$, $\mathbf{s}$ are orthogonal unit vectors along the principal axes of the ellipse of $\mathbf{D}$.

If the bivector $\mathbf{H}$ is defined by $\mathbf{H} = e^{i\phi}\mathbf{D}$ with $\mathbf{D} = \mathbf{a} + i\mathbf{b}$, then

$$\mathbf{H} = (\cos\phi\,\mathbf{a} - \sin\phi\,\mathbf{b}) + i(\sin\phi\,\mathbf{a} + \cos\phi\,\mathbf{b}), \tag{2}$$

so that the ellipse associated with $\mathbf{H}$ is

$$\begin{aligned} \mathbf{r} &= (\cos\phi\,\mathbf{a} - \sin\phi\,\mathbf{b})\cos\theta + (\sin\phi\,\mathbf{a} + \cos\phi\,\mathbf{b})\sin\theta \\ &= \mathbf{a}\cos(\theta - \phi) + \mathbf{b}\sin(\theta - \phi). \end{aligned} \tag{3}$$

This is precisely the ellipse associated with $\mathbf{D}$. Note that here ϕ is a given quantity and θ is a variable. This result is due to MacCullagh, [8], who set its equivalent as an examination question in Trinity College in 1847.

MacCullagh's theorem, described by Hamilton as "a remarkable use of the symbol i," is central to the use of bivectors in the description of wave propagation.

MacCullagh's theorem means that beginning with the pair $(\mathbf{a}, \mathbf{b})$ which defines an ellipse, then $e^{i\phi}(\mathbf{a} + i\mathbf{b})$ gives another pair of conjugate radii of the same ellipse. Now taking $\phi = \omega t$ where ω is a real constant and t denotes time, $e^{i\omega t}(\mathbf{a} + i\mathbf{b})$ gives at any time t a pair of conjugate radii of the ellipse. The tip of the real vector $\cos\omega t\,\mathbf{a} - \sin\omega t\,\mathbf{b}$ moves on the ellipse. Its period of oscillation is $2\pi/\omega$. The pair $(\mathbf{a}, \mathbf{b})$ is rotated, but not rigidly, into another pair of conjugate radii.

Let two bivectors $\mathbf{D}$ and $\mathbf{E}$ be coplanar and orthogonal: $\mathbf{D} \cdot \mathbf{E} = 0$. Now $\mathbf{D}$ may be written

$$\mathbf{D} = e^{i\phi}(\mathbf{a} + i\mathbf{b}) = e^{i\phi}(a\mathbf{i} + ib\mathbf{j}),$$

where $\mathbf{a}$ and $\mathbf{b}$ are along the principal semi-axes of the ellipse of $\mathbf{D}$ and $\mathbf{i}$, $\mathbf{j}$ are unit vectors. Because $\mathbf{E}$ is coplanar with $\mathbf{D}$, it may be written $\mathbf{D} = \alpha\mathbf{i} + \beta\mathbf{j}$ for some scalars α, β. Then $\mathbf{D} \cdot \mathbf{E} = 0$ gives $\alpha a + i\beta b = 0$ so that $\mathbf{E} = \lambda(a\mathbf{j} - ib\mathbf{i})$ for some λ. Hence the major and minor axes of the ellipse of $\mathbf{E}$ are respectively along the minor and major axes of the ellipse of $\mathbf{D}$. Also the ellipses of $\mathbf{E}$ and $\mathbf{D}$ are similar – they have the same aspect ratio (a/b) and they are described in the same sense. Hence Gibbs' result [4] follows: if $\mathbf{D} \cdot \mathbf{E} = 0$ with $\mathbf{D}$ and $\mathbf{E}$ coplanar, the ellipse of $\mathbf{E}$ is similar and similarly situated to the ellipse of $\mathbf{D}$ rotated through a quadrant in its plane. Both ellipses are described in the same sense. (See also [5].)

If two coplanar bivectors $\mathbf{D}$ and $\mathbf{E}$ are such that $\mathbf{D} \cdot \bar{\mathbf{E}} = 0$, then the ellipse of $\mathbf{E}$ is similar and similarly situated to the ellipse of $\mathbf{D}$ when rotated through a quadrant. However the ellipses are described in opposite senses. Then $\mathbf{D}$ and $\mathbf{E}$ are said to be **oppositely polarized**. For example, $\mathbf{D}$ and $\mathbf{E}$ given by

$$\mathbf{D} = \mathbf{i} + i\mathbf{j}, \quad \mathbf{E} = \mathbf{i} - i\mathbf{j}$$

and by

$$\mathbf{D} = \mathbf{i} + im\mathbf{j}, \quad \mathbf{E} = m\mathbf{i} - i\mathbf{j},$$

where $m = \bar{m}$, are oppositely polarized.

Waves

Using rectangular Cartesian coordinate axes $Oxyz$, with the z-axis along the propagation direction, plane transverse homogeneous waves are described by a vector field $\mathbf{u}(z,t)$ of the form

$$\mathbf{u}(z,t) = \{\mathbf{A}e^{i\tau}\}^{+}, \quad \tau = \kappa z - \omega t. \tag{4}$$

Here $\mathbf{A}$ is called the "amplitude bivector" and is constant. It lies in the x-y plane:

$$\mathbf{A} = A_1\mathbf{i} + A_2\mathbf{j}. \tag{5}$$

Also κ and ω are real constants.

In general, (4) describes an infinite train of homogeneous waves, propagating in the z-direction, elliptically polarized in the x-y plane. Its period is $2\pi/(\omega)$ and its wavelength is $2\pi/(\kappa)$.

The intensity I of the wave is defined by

$$I = \mathbf{A}\cdot\bar{\mathbf{A}} = A_1\bar{A}_1 + A_2\bar{A}_2. \tag{6}$$

Writing the complex numbers A_1, A_2 in terms of their moduli and arguments,

$$A_1 = a_1e^{i\delta_1}, \quad A_2 = a_2e^{i\delta_2}, \tag{7}$$

then

$$\mathbf{A} = \mathbf{c} + i\mathbf{d}, \quad \mathbf{c} = a_1\cos\delta_1\mathbf{i} + a_2\cos\delta_2\mathbf{j}, \quad \mathbf{d} = a_1\sin\delta_1\mathbf{i} + a_2\sin\delta_2\mathbf{j}. \tag{8}$$

The components u_1 and u_2 of the vector field $\mathbf{u}(z,t)$ along the x and y axes, respectively, are given by

$$u_1 = a_1\cos(\tau+\delta_1), \quad u_2 = a_2\cos(\tau+\delta_2). \tag{9}$$

The polarization ellipse is contained within a rectangular box whose sides are parallel to the x and y axes, and have lengths $2a_1$ and $2a_2$, respectively. The amplitudes a_1 and a_2, and the phase difference $\delta = \delta_1 - \delta_2$ of the orthogonal field components (9) may be obtained from observations. The polarization ellipse may be constructed if a_1, a_2, δ are known, [3, 5].

Stokes Parameters

Following Stokes, [1], consider the possibility of resolving the given wave (4) into two given elliptically polarized waves of the same period and wavelength. This amounts to asking whether it is possible to decompose $\mathbf{A}$ in the form

$$\mathbf{A} = \lambda\mathbf{C} + \mu\mathbf{D}, \tag{10}$$

where $\mathbf{C}$ and $\mathbf{D}$ are any two given bivectors coplanar with $\mathbf{A}$ such that $\mathbf{C}\cdot\bar{\mathbf{C}} = \mathbf{D}\cdot\bar{\mathbf{D}} = 1$ and λ, μ are to be determined. Now choosing $\mathbf{C}^\star$ and $\mathbf{D}^\star$ so that $\mathbf{C}\cdot\mathbf{C}^\star = \mathbf{D}\cdot\mathbf{D}^\star = 0$, it follows that

$$\lambda\mathbf{C}\cdot\mathbf{D}^\star = \mathbf{A}\cdot\mathbf{D}^\star, \quad \mu\mathbf{D}\cdot\mathbf{C}^\star = \mathbf{A}\cdot\mathbf{C}^\star. \tag{11}$$

Hence λ, μ do not exist if $\mathbf{C}\cdot\mathbf{D}^\star = 0$, or equivalently in this case $\mathbf{D}\cdot\mathbf{C}^\star = 0$. This means that the bivectors $\mathbf{C}$ and $\mathbf{D}$ are parallel: $\mathbf{D} = \gamma\mathbf{C}$ for some γ. Indeed, suppose, without loss, that

$$\mathbf{C} = (\mathbf{i} + im\mathbf{j})/(1+m^2)^{\frac{1}{2}}, \quad \mathbf{D} = (p\mathbf{i} + q\mathbf{j})/(p\bar{p} + q\bar{q})^{\frac{1}{2}}, \tag{12}$$

with $m = \bar{m}$, so that

$$\mathbf{C}^\star = (m\mathbf{i} + i\mathbf{j})/(1+m^2)^{\frac{1}{2}}, \quad \mathbf{D}^\star = (q\mathbf{i} - p\mathbf{j})/(p\bar{p} + q\bar{q})^{\frac{1}{2}}. \tag{13}$$

Then $\mathbf{C}\cdot\mathbf{D}^\star = 0$ leads to $q = ipm$, which leads to $\mathbf{C}^\star\cdot\mathbf{D} = 0$ and $\mathbf{D} = \{p/(p\bar{p})^{\frac{1}{2}}\}\mathbf{C}$.

If $\mathbf{C}$ and $\mathbf{D}$ are not parallel, then λ, μ are determined and using $\mathbf{C}\cdot\bar{\mathbf{C}} = \mathbf{D}\cdot\bar{\mathbf{D}} = 1$, it follows that

$$I = \mathbf{A}\cdot\bar{\mathbf{A}} = \lambda\bar{\lambda} + \mu\bar{\mu} + \lambda\bar{\mu}\mathbf{C}\cdot\bar{\mathbf{D}} + \bar{\lambda}\mu\bar{\mathbf{C}}\cdot\mathbf{D}. \tag{14}$$

Now choose $\mathbf{C}$ and $\mathbf{D}$ so that their ellipses are oppositely polarized: $\mathbf{C}\cdot\bar{\mathbf{D}} = 0$. Then

$$I = \lambda\bar{\lambda} + \mu\bar{\mu}, \tag{15}$$

which is just the sum of the intensities of the component waves.

In the Introduction it was noted that the real vectors $\mathbf{b}$ and $\mathbf{c}$ were chosen to be orthogonal. Here the bivectors $\mathbf{C}$ and $\mathbf{D}$ are chosen so that $\mathbf{C} \cdot \bar{\mathbf{D}} = 0$–the ellipse of $\mathbf{C}$ is similar and similarly situated to that of $\mathbf{D}$ when rotated through a quadrant, but the ellipses of $\mathbf{C}$ and $\mathbf{D}$ are described in opposite senses.

Now if $\mathbf{r}$ and $\mathbf{s}$ are unit vectors along the major and minor axes, respectively, of the ellipse of $\mathbf{C}$, then $\mathbf{C}$ and $\mathbf{D}$ may be written

$$\mathbf{C} = \cos\beta'\mathbf{r} + i\sin\beta'\mathbf{s}\ , \ \ \mathbf{D} = \sin\beta'\mathbf{r} - i\cos\beta'\mathbf{s}. \tag{16}$$

Let χ' be the azimuth of the major axis of the ellipse of $\mathbf{C}$ with respect to the x-axis. Then $\cos\chi' = \mathbf{r}\cdot\mathbf{i} = \mathbf{s}\cdot\mathbf{j}$.

Using (10) and $\mathbf{C} \cdot \bar{\mathbf{D}} = 0$, it follows that

$$\lambda\bar{\lambda} = (\mathbf{A}\cdot\bar{\mathbf{C}})(\bar{\mathbf{A}}\cdot\mathbf{C}), \tag{17}$$

where

$$\mathbf{A}\cdot\bar{\mathbf{C}} = (A_1\mathbf{i} + A_2\mathbf{j})\cdot(\cos\beta'\mathbf{r} - i\sin\beta'\mathbf{s}).$$

Hence

$$2\lambda\bar{\lambda} = I + Q\cos 2\beta'\cos 2\chi' + U\cos 2\beta'\sin 2\chi' + V\sin 2\beta' \tag{18}$$

and similarly

$$2\mu\bar{\mu} = I - Q\cos 2\beta'\cos 2\chi' - U\cos 2\beta'\sin 2\chi' - U\sin 2\beta', \tag{19}$$

where

$$\begin{aligned} I &= A_1\bar{A}_1 + A_2\bar{A}_2, & Q &= A_1\bar{A}_1 - A_2\bar{A}_2, \\ U &= A_1\bar{A}_2 + \bar{A}_1A_2, & V &= i(A_1\bar{A}_2 - \bar{A}_1A_2). \end{aligned} \tag{20}$$

The four quantities I, Q, U, V involve only the components A_1, A_2 of $\mathbf{A}$. They are the Stokes parameters. They all have the same dimensions and are such that

$$I^2 = Q^2 + U^2 + V^2. \tag{21}$$

For every decomposition of the wave with amplitude bivector $\mathbf{A}$ into two oppositely polarized waves, the intensities of the two component waves are given as linear combinations of the four Stokes parameters. There is an infinity of such decompositions–a particular decomposition corresponds to a choice of the polarization form of the bivector $\mathbf{C}$, that is, to a choice of the angles β' and χ'.

If β' and χ' are altered to β'' and χ'' (say), then λ and μ are altered to λ'' and μ'' (say). The intensities $\lambda''\bar{\lambda}''$ and $\mu''\bar{\mu}''$ of the two oppositely polarized waves are given by (18) and (19) respectively, with β', χ' replaced by β'', χ'', **but the coefficients I, Q, U, V remain unchanged.**

If $\tan\beta$ is the aspect ratio of the ellipse of $\mathbf{A}$ and χ is its azimuth, it may be shown, [2, 3, 5], that

$$\begin{aligned} I &= \mathbf{A}\cdot\bar{\mathbf{A}}, & Q &= I\cos 2\beta\cos 2\chi, \\ V &= I\sin 2\beta, & U &= \cos 2\beta\sin 2\chi. \end{aligned}$$

This and (21) immediately suggest representation on a sphere–the Poincaré sphere. But that is another story.

References

[1] G. G. Stokes, *On the composition and resolution of streams of polarized light from different sources.* Trans. Cambridge Phil. Soc. **9** (1852), 399-416.

[2] R. M. A. Azzam and N. M. Bashara, Ellipsometry and Polarised Light. North-Holland: Amsterdam (1977).

[3] M. Born and E. Wolf, Principles of Optics (sixth edition). Pergamon: Oxford (1980).

[4] J. W. Gibbs, Elements of Vector Analysis. Privately printed (1881, 1884). (Published in Scientific Papers, Vol. 2, pp. 17-90. Dover: New York (1961).)

[5] Ph. Boulanger and M. Hayes, Bivectors and Waves in Mechanics and Optics. Chapman and Hall: London (1993).

[6] W. R. Hamilton, Lectures on Quaternions. Hodges and Smith: Dublin (1853).

[7] R. J. Strutt (Fourth Baron Rayleigh), Life of John William Strutt, Third Baron Rayleigh. The University of Wisconsin Press: Madison (1968).

[8] J. MacCullagh, *On total reflexion*, Proc. Royal Irish Acad. **3** (1847), 49-51.

Acknowledgement I am very grateful to T. J. Laffey for his comments.

M. Hayes,
Mathematical Physics Department,
University College,
Belfield,
Dublin 4.

WRITING COMMUTATORS OF GROUP COMMUTATORS AS PRODUCTS OF CUBES

Peter V. Hegarty

Abstract We derive an upper bound for the number of cubes needed to write a commutator of group commutators as a product of cubes.

1. Introduction

If a and b are elements of a group G, we define their commutator $[a,b]$ to be the group element $a^{-1}b^{-1}ab$. It is well known that groups of exponent 3 are metabelian–for a proof see [2], pp. 382-3. Consequently, in the free group F_4 on four generators x, y, z and w, the "commutator of commutators" $[[x,y],[z,w]]$ can be expressed as a product of cubes of elements of F_4. In the survey article [1], R. Lyndon poses the problem of finding such an expression which contains the smallest possible number of cubes. At this point it is instructive to recall, by way of analogy, the simple and well known fact that, in the free group F_2 on two generators x and y, the commutator $[x,y]$ can be written as a product of 3 squares, but of no fewer:

$$[x,y] = (x^{-1})^2(xy^{-1})^2y^2.$$

Lyndon's problem, by contrast, seems to be more difficult. In this note, I will show that $[[x,y],[z,w]]$ can be expressed as a product of 85 cubes. This will, I hope, provide a benchmark for future progress on the problem. Following my proof, one could write down an explicit expression, but in the interests of saving space I shall not do so here.

2. Notation

From now on, all work takes place in the free group F_4 on the four generators x, y, z, w. To simplify the presentation of the proof to