Proceedings of the 9th Solar Polarization Workshop SPW9
A workshop held in Göttingen, Germany, from August 26 to August 30, 2019
Achim Gandorfer, Andreas Lagg, Kerstin Raab (eds.)

DOI: 10.17617/2.3182802

Stokes vectors and Minkowski spacetime: Structural parallels

J.O. Stenflo[1,2,*]

[1]Institute for Particle Physics and Astrophysics, ETH Zurich, CH-8093 Zurich
[2]Istituto Ricerche Solari Locarno (IRSOL), Via Patocchi, CH-6605 Locarno, Switzerland
[*]*Email:* stenflo@astro.phys.ethz.ch

Abstract. Polarized radiative transfer is most effectively described in terms of Stokes 4-vectors and Mueller 4×4 matrices. Here we show that the Stokes formalism has astounding structural parallels with the formalism used for relativity theory in Minkowski spacetime. The structure and symmetry properties of the Mueller matrices are the same as those for the matrix representations of the electromagnetic tensor and the Lorentz transformation operator. Comparison between these various matrices shows that the absorption terms η_k in the Mueller matrix directly correspond to the electric field components E_k in the electromagnetic tensor and the Lorentz boost terms γ_k in the Lorentz transformation matrix, while the anomalous dispersion terms ρ_k correspond to the magnetic field components B_k and the spatial rotation angles ϕ_k. In a Minkowski-type space spanned by the Stokes I, Q, U, V parameters, the Stokes vector for 100 % polarized light is a null vector living on the surface of null cones, like the energy-momentum vector of massless particles in ordinary Minkowski space. Stokes vectors for partially polarized light live inside the null cones like the momentum vectors for massive particles. In this description the depolarization of Stokes vectors appears as a "mass" term, which has its origin in a symmetry breaking caused by the incoherent superposition of uncorrelated fields or wave packets, without the need to refer to a ubiquitous Higgs field as is done in particle physics. The rotational symmetry of Stokes vectors and Mueller matrices is that of spin-2 objects, in contrast to the spin-1 nature of the electromagnetic field. The reason for this difference is that the Stokes objects have substructure: they are formed from bilinear tensor products between spin-1 objects, the Jones vectors and Jones matrices. The governing physics takes place at the substructure level.

1 Introduction

The different disciplines in physics often evolve in isolation from each other, developing their own formalisms and terminologies. This difference in concepts, language, and notations hampers the mutual understanding and deepens the divisions. It has the effect of obscuring the circumstance that the different disciplines may have much more in common than meets the eye, and that much of what seems to be different is of cultural and not of intrinsic origin.

In the present paper we highlight the non-trivial circumstance that the mathematical symmetries and structures that govern the Mueller matrices in polarization theory are the same as those of the electromagnetic tensor in Maxwell theory and the Lorentz transformations in relativistic physics, all of which obey the same Lorentz algebra. Absorption, electric fields, and Lorentz boosts are governed by the symmetric part of the respective transformation matrix, while dispersion, magnetic fields, and spatial rotations are represented by the antisymmetric part. These two aspects can be unified in terms of complex-valued matrices, where the symmetric and antisymmetric aspects represent the real and imaginary parts, respectively.

We can extend this comparison by introducing a Minkowski metric for a 4D space spanned by the four Stokes parameters I, Q, U, V. In this description Stokes vectors that represent 100 % polarization are null vectors, while partial depolarization causes the Stokes vector to lie inside the null cones like the energy-momentum vectors of massive particles in ordinary spacetime. This comparison points to an analogy between depolarization (which can be seen as a symmetry breaking) and the appearance of mass. Another interesting property is that, in contrast to electromagnetism and Lorentz transformations, Stokes vectors and Mueller matrices have the rotational symmetry of spin-2 objects because they have substructure: they are formed from bilinear products of spin-1 objects. Here we try to expose these potentially profound connections and discuss their meanings.

We start in Sect. 2 by showing the relations between the Stokes formalism in polarization physics and the covariant formulation of the Maxwell theory of electromagnetism. After clarifying the symmetries it is shown how the introduction of complex-valued matrices leads to an elegant, unified formulation. In Sect. 3 we show how the Lorentz transformations with its boosts and spatial rotations have the same structure. The introduction of a Minkowski metric for polarization space in Sect. 4 reveals a null cone structure for Stokes vectors that represent fully polarized light. It further brings out an analogy between depolarization caused by the incoherent superposition of fields and the appearance of what may be interpreted as a mass term. In Sect. 5 we highlight the circumstance that the Stokes vectors and Mueller matrices in polarization physics have the symmetry of spin-2 objects, because they have substructure, being formed from bilinear products of vector objects. Section 6 summarizes the conclusions.

2 The electro-magnetic analogy

Let $\boldsymbol{S}_\nu$ be the 4D Stokes vector for frequency ν. Explicitly, in terms of its transposed form (with superscript T) and omitting index ν for clarity of notation, $\boldsymbol{S}^T \equiv (S_0,\, S_1,\, S_2,\, S_3) \equiv (I,\, Q,\, U,\, V)$. The equation for the transfer of polarized radiation can then be written as

$$\frac{\mathrm{d}\boldsymbol{S}_\nu}{\mathrm{d}\tau_c} = (\boldsymbol{\eta} + \boldsymbol{I})\,\boldsymbol{S}_\nu \, - \boldsymbol{j}_\nu/\kappa_c\,, \tag{1}$$

where τ_c is the continuum optical depth, $\boldsymbol{I}$ is the 4×4 identity matrix (representing continuum absorption), κ_c is the continuum absorption coefficient, $\boldsymbol{j}_\nu$ is the emission 4-vector, while the Mueller absorption matrix $\boldsymbol{\eta}$ that represents the polarized processes due to the atomic line transitions is

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_I & \eta_Q & \eta_U & \eta_V \\ \eta_Q & \eta_I & \rho_V & -\rho_U \\ \eta_U & -\rho_V & \eta_I & \rho_Q \\ \eta_V & \rho_U & -\rho_Q & \eta_I \end{pmatrix}. \tag{2}$$

For a detailed account of Stokes vector polarization theory with its notations and terminology we refer to the monographs of Stenflo (1994) and Landi Degl'Innocenti & Landolfi (2004).

While $\eta_{I,Q,U,V}$ represents the absorption terms for the four Stokes parameters I, Q, U, V, the differential phase shifts, generally referred to as anomalous dispersion or magnetooptical effects, are represented by the $\rho_{Q,U,V}$ terms. They are formed, respectively, from the imaginary and the real part of the complex refractive index that is induced when the atomic medium interacts with the radiation field.

We notice that $\boldsymbol{\eta}$ can be expressed as the sum of two matrices: a symmetric matrix that only contains the η terms, and an anti-symmetric matrix that only contains the ρ terms. Antisymmetric matrices represent spatial rotations, as will be particularly clear in Sect. 3 when comparing with Lorentz transformations.

These symmetries turn out to be identical to the symmetries that govern both the Maxwell theory for electromagnetism and the Lorentz transformation of the metric in relativity. Here we will compare these various fields of physics to cast light on the intriguing connections. We begin with electromagnetism.

The analogy between the Stokes formalism and electromagnetism has previously been pointed out by Landi Degl'Innocenti & Landi Degl'Innocenti (1981) using a 3+1 (space + time) formulation rather than the covariant formulation in Minkowski spacetime. It is however only with the covariant formulation that the correspondences become strikingly transparent.

The Lorentz electromagnetic force law when written in covariant 4D form is

$$\frac{\mathrm{d}p^\alpha}{\mathrm{d}\tau} = \frac{e}{m}\, F^{\alpha}_{\ \beta}\, p^\beta \,, \tag{3}$$

where p^α are the components of the contravariant energy-momentum 4-vector, e and m are the electric charge and mass of the particle, and $F^{\alpha}_{\ \beta}$ are the components of the electromagnetic tensor, which has the representation of a 4×4 matrix:

$$F^{\alpha}_{\ \beta} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}. \tag{4}$$

A comparison between Eqs. (2) and (4) immediately reveals the correspondence between absorption η and the electric E field on the one hand and between anomalous dispersion ρ and the magnetic B field on the other hand:

$$\begin{aligned} \eta_{Q,U,V} &\longleftrightarrow E_{x,y,z} \\ \rho_{Q,U,V} &\longleftrightarrow B_{x,y,z}\,. \end{aligned} \tag{5}$$

Let us follow up this structural comparison in terms of a unified description, where absorption and dispersion are combined into a complex-valued absorption. η is proportional to the Voigt function $H(a, v_q)$, with η_0 as the proportionality constant, a the dimensionless damping parameter, and v_q the dimensionless wavelength or frequency parameter. Index q, with $q = 0, \pm 1$, indicates the differential shift of the wavelength scale for atomic transitions with magnetic quantum number $m_{\rm lower} - m_{\rm upper} = q$. Similarly ρ is proportional to $2F(a, v_q)$, where F is the line dispersion function.

$$\begin{aligned} \eta_{I,Q,U,V} &= \ \eta_0\, H_{I,Q,U,V}\,, \\ \rho_{Q,U,V} &= 2\eta_0\, F_{Q,U,V}\,. \end{aligned} \tag{6}$$

In the unified description the H and F functions are combined into the complex-valued

$$\mathcal{H}(a, v_q) \equiv H(a, v_q) - 2i\, F(a, v_q)\,, \tag{7}$$

which now represents the building blocks when forming the corresponding quantities with indices I, Q, U, V to refer to the respective Stokes parameters. $\mathcal{H}_{I,Q,U,V}$ can be combined into the 4-vector

$$\boldsymbol{\mathcal{H}} \equiv \begin{pmatrix} \mathcal{H}_I \\ \mathcal{H}_Q \\ \mathcal{H}_U \\ \mathcal{H}_V \end{pmatrix} \equiv \begin{pmatrix} \mathcal{H}_0 \\ \mathcal{H}_1 \\ \mathcal{H}_2 \\ \mathcal{H}_3 \end{pmatrix} . \tag{8}$$

From the above follows how $\mathcal{H}_k$ is related to and unifies the corresponding absorption and dispersion parameters η_k and ρ_k:

$$\eta_k - i\,\rho_k = \eta_0\,\mathcal{H}_k\,. \tag{9}$$

Let us next define the three symmetric matrices $\boldsymbol{K}^{(k)}$ and three antisymmetric matrices $\boldsymbol{J}^{(k)}$ through

$$\begin{aligned} \boldsymbol{K}^{(k)}_{0j} &\equiv \boldsymbol{K}^{(k)}_{j0} \equiv 1\,, \\ \boldsymbol{J}^{(k)}_{ij} &\equiv -\boldsymbol{J}^{(k)}_{ji} \equiv -\,\varepsilon_{ijk}\,, \end{aligned} \tag{10}$$

where ε_{ijk} is the Levi-Civita antisymmetric symbol. We further define the complex-valued matrix

$$\boldsymbol{T}^{(k)} \equiv \boldsymbol{K}^{(k)} - i\,\boldsymbol{J}^{(k)}\,. \tag{11}$$

Then the Mueller matrix from Eq. (2) becomes

$$\boldsymbol{\eta} - \eta_I \boldsymbol{I} = \eta_0\,\mathrm{Re}\,(\,\mathcal{H}_k\,\boldsymbol{T}^{(k)}\,)\,. \tag{12}$$

Let us similarly define the complex electromagnetic vector

$$\boldsymbol{\mathcal{E}} \equiv \boldsymbol{E} - i\,\boldsymbol{B}\,, \tag{13}$$

which in quantum mechanics represents photons with positive helicity. Then the electromagnetic tensor $F^{\alpha}_{\ \beta}$ of Eq. (4) can be written as

$$\boldsymbol{F} = \mathrm{Re}\,(\,\mathcal{E}_k\,\boldsymbol{T}^{(k)}\,)\,. \tag{14}$$

Comparison with Eq. (12) again brings out the structural correspondence between the Mueller matrix and the electromagnetic tensor, this time in a more concise and compact form. It also shows how the electric and magnetic fields are inseparably linked, as the real and imaginary parts of the same complex vector.

It may be argued that the structural similarity between the Mueller and electromagnetic formalisms is not unexpected, since the underlying physics that governs the Mueller matrices is the electromagnetic interactions between matter and radiation. The atomic transitions are induced by the oscillating electromagnetic force of the ambient radiation field when it interacts with the atomic electrons, and this interaction is governed (in the classical description) by the force law of Eq. (3) with its electromagnetic tensor. This is however not the whole story, since there are also profound differences: As we will see in Sect. 5 Stokes vectors and Mueller matrices behave like spin-2 objects, while the electromagnetic tensor is a spin-1 object. Another interesting aspect in the comparison between Stokes vectors and the energy-momentum 4-vector is that depolarization of Stokes vectors acts as if the corresponding 4-vector has acquired "mass", as will be shown in Sect. 4. Before we turn to these topics we will in the next section show the correspondence between the Mueller matrix and the Lorentz transformation matrix.

3 Lorentz transformations and the Mueller absorption matrix

Let $\boldsymbol{X}$ be the spacetime 4-vector:

$$\boldsymbol{X} \equiv \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{15}$$

With the Lorentz transformation $\boldsymbol{\Lambda}$ we transfer to a new system $\boldsymbol{X}'$:

$$\boldsymbol{X}' = \boldsymbol{\Lambda}\,\boldsymbol{X}\,. \tag{16}$$

$\boldsymbol{\Lambda}$ represents rotations in Minkowski space, composed of three spatial rotations ϕ_k and three boosts γ_k, which may be regarded as imaginary rotations. Let us combine them into complex rotation parameters α_k through

$$\alpha_k \;\equiv\; \gamma_k + i\,\phi_k\,. \tag{17}$$

Then the Lorentz transformation $\boldsymbol{\Lambda}$ can be written as

$$\boldsymbol{\Lambda} \equiv e^{\boldsymbol{V}}\,, \tag{18}$$

where

$$\boldsymbol{V} \;=\; \mathrm{Re}\,(\,\alpha_k\,\boldsymbol{T}^{(k)}\,) \;=\gamma_k\,\boldsymbol{K}^{(k)} +\; \phi_k\,\boldsymbol{J}^{(k)}\,. \tag{19}$$

Explicitly,

$$\boldsymbol{V} \;=\begin{pmatrix} 0 & \gamma_x & \gamma_y & \gamma_z \\ \gamma_x & 0 & -\phi_z & \phi_y \\ \gamma_y & \phi_z & 0 & -\phi_x \\ \gamma_z & -\phi_y & \phi_x & 0 \end{pmatrix}. \tag{20}$$

Note that in quantum field theory a convention for the definition of the $\boldsymbol{K}$ and $\boldsymbol{J}$ matrices that define the Lorentz algebra is used, which differs by the factor of the imaginary unit i from the convention of Eq. (10) used here, in order to make $\boldsymbol{K}$ anti-hermitian and $\boldsymbol{J}$ hermitian (Zee 2010).

Comparing with the Mueller matrix and the electromagnetic tensor we see the correspondence

$$\begin{aligned} \eta_k \;&\longleftrightarrow\;\; \gamma_k \;\longleftrightarrow\;\; E_k \\ \rho_k \;&\longleftrightarrow -\phi_k \;\longleftrightarrow\;\; B_k\,. \end{aligned} \tag{21}$$

The minus sign in front of ϕ_k is only due to the convention adopted for defining the sense of rotations and is therefore irrelevant for the following discussion of the physical meaning of the structural correspondence.

Relations (21) show how the Lorentz boosts γ_k, which change the energy and momentum of the boosted object, relate to both absorption η_k and electric field E_k, while the spatial rotations ϕ_k, which do not affect the energy but change the phase of the rotated object, relate to both dispersion or phase shift effects ρ_k and to magnetic fields B_k.

4 Analogy between depolarization and the emergence of mass

The structural correspondence between the 4×4 Mueller absorption matrix, the matrix representation of the covariant electromagnetic tensor, and the Lorentz transformation matrix suggests that there may be a deeper analogy or connection between the 4D Stokes vector space and 4D spacetime. Let us therefore see what happens when we introduce the Minkowski metric to the Stokes vector formalism. The usual notation for the Minkowski metric is $\eta_{\mu\nu}$, but to avoid confusion with the absorption matrix $\boldsymbol{\eta}$ that we have been referring to in the present paper, we will use the notation $\boldsymbol{g}$ or $g_{\mu\nu}$ that is generally reserved for a general metric, but here we implicitly assume that we are only dealing with inertial frames, in which $g_{\mu\nu} = \eta_{\mu\nu}$.

Assume that $\boldsymbol{I}_\nu = \boldsymbol{S}_\nu$ is the 4D Stokes vector, with its transpose being $\boldsymbol{I}_\nu^T = (I_\nu,\, Q_\nu,\, U_\nu,\, V_\nu)$. The scalar product in Minkowski space is then

$$\boldsymbol{I}_\nu^T \boldsymbol{g}\, \boldsymbol{I}_\nu \,=\, I_\nu^2 \,-\, (\,Q_\nu^2 \,+\, U_\nu^2 \,+\, V_\nu^2\,)\,, \tag{22}$$

which also represents the squared length of the Stokes vector in Minkowski space.

We know from polarization physics that the right-hand-side of Eq. (22) is always ≥ 0, and equals zero only when the light beam is 100 % (elliptically) polarized. Such fully polarized, pure or coherent states are thus represented by null vectors, in exactly the same way as the energy-momentum 4-vector $\boldsymbol{p}$ of massless particles are also null vectors on the surface of null cones. The energy-momentum vectors of massive particles live inside the null cones. Similarly the Stokes vectors live inside and not on the surface of null cones only if the light is not fully but partially polarized.

This comparison raises the question whether there is some deeper connection between depolarization and the appearance of mass. In polarization physics all individual (coherent) wave packages are 100 % polarized, and any coherent superposition of such wave packages is also fully polarized. Partial polarization occurs exclusively as a result of the *incoherent* superposition of different, uncorrelated wave packages. In such cases it is customary to represent the intensity I_ν, which represents the energy or the number of photons carried by the beam, as consisting of two parts, one fraction p_ν that is fully polarized, and one fraction with intensity $I_{\nu,u}$ that is unpolarized, with transposed Stokes vector $I_{\nu,u}\,(1,\,0,\,0,\,0)$:

$$I = p_\nu\, I \,+\, I_u\,, \tag{23}$$

where we have omitted index ν for simplicity except for p_ν (to distinguish it from the momentum vector p below). This fractional polarization p_ν is

$$p_\nu = \frac{I - I_u}{I} = \frac{(Q^2 + U^2 + V^2)^{1/2}}{I}\,. \tag{24}$$

In comparison, in particle physics, the scalar product for the 4D energy momentum vector $\boldsymbol{p}$ is

$$\boldsymbol{p}^T \boldsymbol{g}\, \boldsymbol{p} \,=\, m^2\, c^2\,, \tag{25}$$

from which the well-known Dirac equation

$$E^2 = p^2\, c^2 \,+\, m^2\, c^4 \tag{26}$$

follows. While the emergence of the mass term corresponds to the emergence of the unpolarized component I_u, Eqs. (23) and (26) look different, because the decomposition in Eq. (23) has been done for the unsquared intensity I, while in Eq. (26) it is in terms of the squared components. Since we have the freedom to choose different ways to mathematically decompose a quantity, this difference is not of particular physical significance.

In current quantum field theories (QFT) the emergence of mass requires the spontaneous breaking of the gauge symmetry, for which the Higgs mechanism has been invented. It is postulated that all of space is permeated by a ubiquitous Higgs field, which when interacting with the field of a massless particle breaks the symmetry. When the particle gets moved to the non-symmetric state it acquires mass. Because the phases of the Higgs field and the field of the initially massless particle are uncorrelated, the superposition of the fields is incoherent, which may be seen as one reason for the breaking of the symmetry.

In polarization physics the emergence of depolarization may also be interpreted as a symmetry breaking, caused by the incoherent superposition of different wave fields. Incoherence means that the phases of the superposed fields are uncorrelated, which has the result that the interference terms, all of which are needed to retain the symmetry, vanish.

5 Stokes vectors as spin-2 objects

An object with spin s varies with angle of rotation θ as $s\,\theta$. For $s = \frac{1}{2}$ one has to rotate 4π radians to return to the original state, for $s = 2$ one only needs to rotate π radians, and so on. Ordinary vectors, like the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$, rotate like spin-1 objects. It may therefore come as a surprise that the Stokes vector rotates with twice the angle, like a spin-2 object, in spite of the identical symmetry properties of the Mueller matrix and the electromagnetic tensor.

The resolution to this apparent paradox is found by distinguishing between the kind of spaces in which the rotations are performed. In the Minkowski-type space that is spanned by I, Q, U, V as coordinates, which is the Poincaré space in polarization physics for a fixed and normalized intensity I, the transformation properties are indeed those of a real vector, a spin-1 object. However, besides Poincaré space the Stokes vector also lives in ordinary space, and a rotation by θ of a vector in Poincaré space corresponds to a rotation in ordinary space by 2θ. While being a spin-1 object in Poincaré space, the same object becomes a spin-2 object in ordinary space.

The reason why it becomes a spin-2 object is that the Stokes vector has substructure: it is formed from tensor products of Jones vectors. Similarly Mueller matrices for coherent (100 % polarized) wave packages are formed from tensor products of Jones matrices. While the Jones vectors and matrices are spin-1 objects in ordinary space, the bilinear products between them become spin-2 objects.

The fundamental physics that governs the polarization physics does not manifest itself at the level of these spin-2 objects, because the basic processes are the electromagnetic interactions between the radiation field and the electrons (which may be bound in atoms), and these interactions are described at the spin-1 level (since the electromagnetic waves represent a spin-1 vector field). The Jones matrices, or, in QM terminology, the Kramers-Heisenberg scattering amplitudes, contain the fundamental physics. They are the basic building blocks for the bilinear products, the spin-2 objects.

This discussion points to the possibility that the physics of other types of spin-2 objects, like the metric field in general relativity, may be hidden, because the governing physics may take place within a spin-1 substructure level and would remain invisible if the spin-2 field would be (incorrectly) perceived as fundamental, without substructure.

6 Conclusions

Comparison between the Stokes formalism, the covariant formulation of electromagnetism, and the Lorentz transformation shows that they all share the same Lie algebra, namely the algebra of the Lorentz group. This algebra is 6-dimensional (for instance in the case of electromagnetism we have three electric field components + three magnetic field components). While this is the algebra that is known to govern Lorentz transformations and the related covariant formulation of electromagnetism, it is not obvious why this group algebra should also apply to the transformation of Stokes 4-vectors, which have been constructed with the aim of being a powerful tool for the treatment of partially polarized light.

In spite of the common underlying group structure, there is also a profound difference. While the electromagnetic field vectors and tensors are objects of a vector space with rotational properties of spin-1 objects, Stokes vectors and Mueller matrices have the rotational symmetries of spin-2 objects, because they are formed from tensor products of spin-1 objects. This vector-field substructure contains the governing physics, where everything is coherent, and where in the quantum description the probability amplitudes or wave functions live and get linearly superposed to form mixed states with certain phase relations. When we go to the spin-2 level by forming bilinear products between the probability amplitudes, which generates observable probabilities, or when we form bilinear products between electric field vectors to generate quantities that represent energies or photon numbers, we get statistical quantities (probabilities or energy packets) over which we can form ensemble averages. If the phase relations of the mixed states in the substructure are definite, we get interference effects and 100 % polarization for the ensemble averages, while if the phase relations contain randomness (incoherent superposition) we get partial polarization.

When comparing the Stokes 4-vector with the energy-momentum 4-vector of a particle, 100 % elliptically polarized light corresponds to massless particles, while the Stokes vector for partially polarized light corresponds to the energy-momentum vector of a massive particle. Depolarization thus has an effect as if the Stokes vector has acquired “mass” by a symmetry breaking that is caused by the destruction of coherences between mixed states.

References

Landi Degl’Innocenti, E., & Landi Degl’Innocenti, M. 1981, Nuovo Cimento, 62, B1

Landi Degl’Innocenti, E., & Landolfi, M. 2004, Polarization in Spectral Lines, vol. 307 of Astrophysics and Space Science Library (Kluwer)

Stenflo, J. O. 1994, Solar Magnetic Fields — Polarized Radiation Diagnostics (Kluwer)

Zee, A. 2010, Quantum Field Theory in a Nutshell: Second Edition (Princeton University Press)