Rotation and Lorentz Transformations in $2 \times 2$ and $4 \times 4$ Complex Matrices and in Polarized-Light Physics

Laurence J. November*

National Solar Observatory
National Optimal Astronomy Observatories†
Sacramento Peak
Sunspot, New Mexico 88349

Submitted by Paul Van Dooren

ABSTRACT

Hermitian $2 \times 2$ matrices exhibit basic 3D rotational and 4D Lorentz transformation properties. These matrices arise naturally in representations of the time-averaged pair products or intensities of any two-element wave, giving rise to the light Stokes-parameter transformation properties on the Poincaré sphere. Equivalent transformations are obtained for $4 \times 4$ anticommuting Hermitian Dirac matrices with two types of unitary matrices, corresponding to rotation and Lorentz transformations. Using exponential matrix representations, the $4 \times 4$ form can be related to the $2 \times 2$ form. The $4 \times 4$ representation has physical significance for the subset of intensity-distinguishable two-element standing-wave modes of a cavity, e.g. light standing waves. There is a basic resemblance between (1) the temporal differential equation for two-element standing waves in time, three observable "Stokes" parameters, and frequency and (2) the Dirac equation for spin-1/2 free-space particle states in time, three momenta, and particle rest mass. This resemblance is the basis for an optical analog with relativistic quantum mechanics which we describe. © Elsevier Science Inc., 1997

* E-mail: lnovember@sunspot.noao.edu.
† NOAO is operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under cooperative agreement with the National Science Foundation.
1. INTRODUCTION

The commutation and anticommutation properties of $2 \times 2$ Hermitian matrices give them natural 3D rotational or 4D Lorentz-transformation properties. The traceless $2 \times 2$ Hermitian matrix squared is a scalar times the identity in the three real orthogonal matrix elements: $\alpha s_1^2 + s_2^2 + s_3^2$. Similarity transformations of $2 \times 2$ Hermitian matrices preserve the matrix determinant, the four-element "Lorentz-transformation" modulus in the four real orthogonal matrix elements: $\alpha s_0^2 - s_1^2 - s_2^2 - s_3^2$. As light polarized intensities are represented in a $2 \times 2$ Hermitian "coherency" matrix, light polarization transformation have associated natural 3D and 4D geometries.

Poincaré's 1892 book *La Lumière* [20, Chapter XII, "Polarisation Rotatoire—Théorie de M. Mallard," pp. 275–301] reported the geometric relations for optical transformations of the light Stokes parameters $s_0$, $s_1$, $s_2$, and $s_3$. The total intensity is $s_0$. The Stokes parameter $s_1$ represents the intensity difference measured between linear polarizers at $0^\circ$ and $90^\circ$, and $s_2$ between linear polarizers at $45^\circ$ and $135^\circ$. The Stokes parameter $s_3$ is the circularly polarized intensity, the intensity difference between right and left circular analyzers. Poincaré pointed out an energy-conserving unitary linear transformation of the two-element light electromagnetic wave vector can be represented as a rotation in the 3D space of the three polarized light intensities. His idea lets us visualize the polarization action of optical retarders as rotations of the Stokes parameters around an axis in $(s_1, s_2, s_3)$ on the "Poincaré sphere."

The reason that the Stokes parameters are restricted to the Poincaré sphere in unitary transformations follows directly from the definition for the light intensities as the time-averaged pair products that can be formed between the two complex extended electromagnetic components representing the light wave. The real electric-field vector for the light wave at a point in space is denoted $E_r(t) = (E_x(t), E_y(t))^T$, where $E_x(t)$ and $E_y(t)$ are the perpendicular electric-field components along $x$ and $y$, both perpendicular to the light travel direction $z$; the superscript $T$ is used to denote transpose for writing the vertical column vector in line. The four nonvanishing time-averaged product pairs, or intensities, are most easily defined using the coherency matrix $J$ for the Stokes vector $s = (s_0, s_1, s_2, s_3)^T$ [6, §10.1-4, pp. 491–508]:

$$J = \left< \tilde{E}_r(t) \otimes \tilde{E}_l(t)^\dagger \right>_t = \frac{1}{2} \begin{pmatrix} s_0 + s_1 & s_2 - is_3 \\ s_2 + is_3 & s_0 - s_1 \end{pmatrix}.$$  \hspace{1cm} (1.1)

where: $\tilde{E}_r(t)$ represents the analytic signal from the real $E_r(t)$, that is, the real function plus $i$ times the Hilbert-transform convolution of the real
function in time \([8, \text{pp. 267–272}]\), \(\otimes\) is the Kronecker outer product; \(\langle \cdot \rangle_t\) denotes the time average. Many useful theorems have been proven for the coherency matrix. Most important is the property that the coherency matrix is a complete descriptor for a monochromatic wave, losing only the absolute wave phase, and is linear in a quasimonochromatic superposition of waves \([22, 2, 6]\). Most physical processes respond exclusively to the intensities, and only processes tuned to the wave frequencies can infer other physical information.

The time-independent \(2 \times 2\) complex Jones matrix \(N\), which transforms the two-element complex electromagnetic wave vector with the linear transformation

\[
\tilde{E}'(t) = N\tilde{E}(t).
\]

transforms the corresponding coherency matrix with the congruency transformation:

\[
J' = NJN^+. \tag{1.3}
\]

The coherency matrix transformation follows simply from the definition: \(J' = \langle \tilde{E}'(t) \otimes \tilde{E}'(t)^\dagger \rangle_t\), where \(N^\dagger\) denotes the transpose conjugate of \(N\).

The 3D rotation and 4D Lorentz groups common to transformations of \(2 \times 2\) Hermitian matrices and to the spatial-temporal transformations of special relativity have been studied extensively by many authors \([21; 13, \text{Section 7-4}; 10]\). Transformations with \(N\) unitary preserve the two eigenvalues of the input-state coherency matrix \(J\), thus conserving the intensity \(s_0\) and the total polarized intensity \((s_1^2 + s_2^2 + s_3^2)^{1/2}\). The unitary \(N\) produces a rotation in the space of \((s_1, s_2, s_3)\) with its eigenvectors corresponding to oppositely directed vectors that define the axis of the rotation in \((s_1, s_2, s_3)\), its eigenvalues being phase factors whose phase difference defines the rotation angle.

The unimodular \(N\) preserves the total determinant of the input coherency matrix, \(\det J = (s_0^2 - s_1^2 - s_2^2 - s_3^2)/4\). This transformation corresponds to a Lorentz boost in \(s_0\) and \((s_1, s_2, s_3)\) \([9]\). By singular-value decomposition of the unimodular \(N\) into the product of a unitary matrix, a diagonal matrix of real values, and another unitary matrix, we understand its physical effect as a rotation, followed by a “Lorentz skewing” along the \(s_1\) axis, followed by a second rotation. In the past we have made use of basic linear-algebraic methods in practical studies of polarized light and in measurement of polarized-light operators, e.g. serial matrix multiplication to represent multiple-optical-element systems, and eigenvalue and singular-value decom-
position methods [19, 18]. The main result of this paper is that the rotation and Lorentz transformations take different forms in multielement standing-wave systems, more akin to the form that appears in relativistic quantum mechanics.

2. 4 × 4 MATRIX ROTATION AND LORENTZ TRANSFORMATIONS

The Lorentz transformation occurs in relativistic quantum mechanics as a 4 × 4 unitary similarity transformation which acts to preserve the four-element relativistic modulus in the Dirac Hamiltonian: 

\[ m^2 = E^2 - p_1^2 - p_2^2 - p_3^2, \]

where \( m \) is the constant particle rest mass, \( E \) the particle energy, and \( (p_1, p_2, p_3) \) the momentum defined by the spatial derivative with respect to the proper time, all written in common units [11]. Dirac could obtain a viable matrix representation for the relativistic Hamiltonian linear in the momenta and particle rest mass, as is required to obtain the correct mass continuity equation using 4 × 4 anticommuting Hermitian matrices.

With 2 × 2 matrices there are three anticommuting Hermitian matrices, represented most conveniently by the Pauli spin matrices \( \sigma_1, \sigma_2, \) and \( \sigma_3 \):

\[
\begin{align*}
\sigma_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, & \sigma_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{align*}
\]

The anticommutation property is easily demonstrated: \( \sigma_j \cdot \sigma_k = -\sigma_k \cdot \sigma_j \) for all \( j, k = 1, 2, 3 \), \( j \neq k \), and \( \sigma_j \cdot \sigma_j = \sigma_0 \) for all \( j \). Thus the square of the traceless 2 × 2 Hermitian matrix \( H_2(z) = z_1 \sigma_1 + z_2 \sigma_2 + z_3 \sigma_3 \) is the vector modulus times the identity: \( H_2(z) \cdot H_2(z) = (z_1^2 + z_2^2 + z_3^2) \sigma_0 = |z|_3 \sigma_0 \); we use the subscript on the modulus of a real vector, like \( |z|_3 \), to remind us of the vector dimension. With the 2 × 2 unitary transformation

\[
H_2(z') = R(\zeta)H_2(z)R(-\zeta),
\]

the vector modulus is preserved: \( |z'|_3^2 = |z|_3^2 \), as can be demonstrated by squaring both sides of the equation. This property can also be proved for more general transformations by unimodular matrices using the determinant relation: \( \det H_2(z) = -z_1^2 - z_2^2 - z_3^2 = -|z|_3^2 \). The general 2 × 2 unitary "rotation" matrix \( R(\zeta) \), which preserves the Hermitian form of \( H_2(z) \), is a function of three real angular parameters, \( \zeta = (\zeta_1, \zeta_2, \zeta_3)^T \), which are the
three real elements in the $2 \times 2$ traceless Hermitian matrix $K(\xi)$, where

$$R(\xi) = \exp[-iK(\xi)].$$

(2.3)

With $4 \times 4$ matrices there are five anticommuting Hermitian matrices, which can be defined in different ways as outer-product combinations of Pauli spin matrices, e.g.

$$\alpha_0 = \sigma_0 \otimes \sigma_0, \quad \alpha_j = \sigma_3 \otimes \sigma_j, \quad \text{for} \quad j = 1, 2, 3,$$

$$\alpha_4 = \sigma_1 \otimes \sigma_0, \quad \alpha_5 = \sigma_2 \otimes \sigma_0; \quad (2.4)$$

then $\alpha_j \cdot \alpha_k = -\alpha_k \cdot \alpha_j$ for all $j, k = 1, \ldots, 5$, $j \neq k$, and $\alpha_j \cdot \alpha_j = \alpha_0$ for all $j$. The anticommutation property is easily verified, since the matrix product distributes under the outer product, e.g. $\alpha_1 \cdot \alpha_2 = (\sigma_2 \otimes \sigma_1) \cdot (\sigma_2 \otimes \sigma_2) = (\sigma_2 \cdot \sigma_2) \otimes (\sigma_1 \cdot \sigma_2)$. Thus the square of the special Hermitian matrix

$$H_4(z) = \sum_{j=1}^{5} z_j \alpha_j$$

(2.5)

is the vector modulus times the identity: $H_4(z) \cdot H_4(z) = \sum_{j=1}^{5} z_j^2 \alpha_0 = |z|^2 \alpha_0$. It follows that unitary similarity transformations of $H_4(z)$ preserve the vector modulus $|z|^2$, and thus can be visualized as rotations in the five-dimensional space of $z$. These properties of anticommuting matrices generalize to arbitrary $2^n$ order, since we add two anticommuting matrices for each doubling of the order, from three spin matrices for order 2, to five spin matrices for order 4, to seven spin matrices for order 8, etc.

Dirac [11] recognized the utility of the $4 \times 4$ anticommuting Hermitian spin matrices for forming a 4D basis with components exhibiting Lorentz-transform invariance by the action of unitary matrices. His three alpha and one beta matrices correspond to $(\alpha_1, \alpha_2, \alpha_3)$ and $\alpha_4$ defined in Equation (2.4), respectively, being related strictly by a similarity transformation corresponding to a change of basis. For the physics, the choice of basis is arbitrary, as we describe in Section 5; the one used here is convenient because it is the natural one for the polarized-light optical analog. The anticommutation property of the elements of $H_4$ lets us define a $4 \times 4$ Hamiltonian matrix linear in the three momenta and rest mass $(p_1, p_2, p_3, m)$, whose square is the scalar $p_1^2 + p_2^2 + p_3^2 + m^2$ times the identity. Thus the operator based upon $H_4$ from Equation (2.5) is a valid Hamiltonian operator providing the
proper Lorentz-transform invariance and whose eigenvalues represent the
allowed energies of the wave function. Curiously the physical system has only
four components, and so needs only four anticommuting spin matrices. While
the $4 \times 4$ representation is required as a minimum, the full set of available
anticommuting matrices is not required, prompting Eddington's suggestion
that this form lacked a certain mathematical symmetry [12].

Subsets of the general five-parameter set in $H_4(z)$ can be generated by
certain unitary matrices. The three-parameter subset $(z_1, z_2, z_3)$ is trans-
formed by the $4 \times 4$ rotation matrix

$$U^{(0)}(\xi) = \exp[-i\sigma_0 \otimes K(\xi)], \quad (2.6)$$

where $K(\xi)$ is a general $2 \times 2$ traceless Hermitian matrix defined by the
three parameters in $\xi$. The similarity transformation of $H_4(z)$ by $U^{(0)}(\xi)$
rotates the three-element vector $(z_1, z_2, z_3)$, preserving the vector modulus
and leaving the other components $z_4$ and $z_5$ unchanged. We think of $U^{(0)}(\xi)$
therefore as producing a rotation in the subset of the parameters that are
coefficients to the three spin matrices $(\alpha_1, \alpha_2, \alpha_3)$, corresponding to the
coefficients for the three quantum-mechanical momenta in the Dirac Hamil-
tonian. The matrices $U^{(0)}(\xi)$ form a closed set under matrix multiplication.

A four-parameter subset of $H_4(z)$ is generated by a different $4 \times 4$
unitary matrix:

$$U^{(1)}(\xi) = \exp[-i\sigma_1 \otimes K(\xi)], \quad (2.7)$$

with $K(\xi)$ a general $2 \times 2$ traceless Hermitian matrix. Then the similarity
transformation generates a four-parameter subset $(z_1, z_2, z_3, z_4)$:

$$H_4(z) = U^{(1)}(\xi) \left( \begin{array}{cc} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{array} \right) U^{(1)}(-\xi) - \sum_{j=1}^{4} z_j \alpha_j, \quad (2.8)$$

where: $z_j = \xi_j \sin(2 |\xi_3|)/|\xi_3|$ for $j = 1, 2, 3$, and $z_4 = \cos(2 |\xi_3|); \mathbf{0}$ denotes
a vector or matrix of zeros of implied order. This solution $z$ is strictly correct
as long as the determinant relation applies: $|\det K(\xi)| \leq 1$. Otherwise a
different $z$ is obtained, based upon a different set of parameters than those
contained in the original $U^{(1)}(\xi)$, reflecting the cyclic form of $U^{(1)}(\xi)$. Thus
the unitary $U_1$ is a Lorentz transformation matrix in that it can be used to
generate the four spin matrices $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ in $H_4(z)$, corresponding to
the three momenta and rest-mass basis matrices in the Dirac Hamiltonian.
Unfortunately $U_1$ does not remain closed under multiplication, so it can be
used to produce a Lorentz transformation by the similarity transform of $H_4(z)$ only with an appropriate matrix-reduction procedure, e.g. [16, Chapter XX, §11, pp. 900–904].

The two special unitary transformation matrices, $U^{(\phi)}(\zeta)$, which transform and generate the four-element subset of $H_4(z)$ relevant for quantum mechanics, exhibit a convenient commutation relation with the $2 \times 2$ rotation matrix $R = \exp[-iK(\zeta)]$:

$$R(\zeta)(\sigma_0 \ 0) = (\sigma_0 \ 0)U^{(0)}(\zeta),$$

(2.9)

with multiplication by the $2 \times 4$ reducing matrix: $(\sigma_0, \sigma_0) = (1, 1) \otimes \sigma_0$. This property is easily verified using the identities

$$U^{(0)}(\zeta) = \sigma_0 \otimes \exp[-iK(\zeta)]$$

(2.10)

and

$$U^{(1)}(\zeta) = \sigma_0 \otimes \cos K(\zeta) - i\sigma_1 \otimes \sin K(\zeta),$$

(2.11)

where the sine and cosine of Hermitian matrices are also Hermitian. The expansion into sine and cosine matrices can be proven for the outer-product form by power-series or eigenvalue decomposition.

Later, in Section 4, we will find this same commutation relation in reducing the standing-light-wave modes of a cavity to an "intensity-distinguishable" subset. By intensity-distinguishable, we mean that the subset of eigenmodes can be distinguished on the basis of measurements of time-averaged pair products of the solution vector elements. Thus we discover that the intensity-distinguishable two-element standing-wave solutions obey a differential equation that resembles the Dirac equation of relativistic quantum mechanics. This resemblance is the basis for an optical analog with relativistic quantum mechanics, which we describe in qualitative terms in Section 5.

3. DIFFERENTIAL EQUATION FOR MULTIELEMENT STANDING WAVES

A standing light wave in a cavity containing a nonconducting medium without charge sources, currents, or material magnetization effects satisfies the differential equation in its electric- and magnetic-field components from
Maxwell's equations, written in Gaussian units:

\[
\left( \nabla^2 - \frac{n^2}{c^2} \frac{\partial^2}{\partial t^2} \right) \tilde{E}_{(\pm)m}(x,t) = \frac{4\pi}{c^2} \frac{\partial^2}{\partial t^2} \tilde{P}(x,t),
\]

where \( n \) is the effective index of refraction for the medium, and \( P(x,t) \) is the nonlinear polarization representing the nonlinear susceptibilities of the resonator medium. We have made the usual assumption that the spatial envelope of the eigenfunction is slowly varying in order to eliminate the term which arises in the \( \nabla_x \times \nabla_x \times E_{(\pm)m}(x,t) \) expansion with a nonzero nonlinear polarization \([7, \S 2.1, pp. 57-62]\). The equations have been transformed in the usual way by applying the Hilbert analytic signal as a linear operator in time to give the complex-extended form for the electromagnetic fields and nonlinear polarization as indicated by the tilde. The analytical extension maps \( \cos \omega t \) into \( \exp(i\omega t) \) and \( \sin \omega t \) into \( \exp(i\omega t - i\pi/2) \), and permits us to treat the separated temporal and spatial parts of the eigenfunctions both conveniently as complex quantities. Of course, the recovery of the equivalent real signal can always be obtained by the identity \( E_{(\pm)m}(x,t) = \frac{1}{2}[E_{(\pm)m}(x,t) + \tilde{E}_{(\pm)m}(x,t)^*] \).

The single index \( m \) is used to represent a discrete counter, or multiple discrete counters for all of the standing-wave solutions in a general finite-sized geometry with boundary conditions. The index \((\pm)\) is used to indicate the two independent spatial eigenfunctions, which are both possible with many geometries as we discuss. In this section and the next we restrict our attention to the homogeneous standing-wave solutions, dropping the nonlinear polarization: \( P(x,t) \to 0 \). While the inclusion of perturbing nonlinear susceptibilities is implicit in this analysis, it is discussed only in qualitative terms in Section 5.

The driving equation (3.1) is written for just two electric-field elements, since the six field elements contained in the two three-vectors for the electric and magnetic fields can be defined with any independent pair \([15, \S 7.1, pp. 269-273; \S 8.2, pp. 339-344]\). It is appropriate to take the electric-field elements \( E_{(\pm)m}(x,t) \) counted by \( j \) in each \( \pm \) spatial eigenfunction and in each mode \( m \) as the two perpendicular components perpendicular to the wavenumber vector. Even though the wavenumber direction can vary from mode to mode or as a function of spatial position in a single mode due to refractive effects with index variations, for many geometric arrangements and boundary conditions it is naturally defined. As this work presents a physical idealization without trying to identify a specific realizable geometric example, we take the third-component direction to be the wavenumber direction for the unperturbed modes.
The eigenmode vector separates into independent spatial and temporal eigenfunctions:

\[
\mathbf{E}_{(\pm)m}(\mathbf{x}, t) = \begin{pmatrix} E_x(\pm)m(\mathbf{x}, t) \\ E_y(\pm)m(\mathbf{x}, t) \end{pmatrix} = \begin{pmatrix} \varphi_x(\pm)m(t)h_x(\pm)m(\mathbf{x}) \\ \varphi_y(\pm)m(t)h_y(\pm)m(\mathbf{x}) \end{pmatrix}.
\]  

(3.2)

The two-element spatial eigenfunctions satisfy the differential equation

\[
\left(\nabla_x^2 + \frac{n^2}{c^2} \omega_m^2\right) \mathbf{h}_{(\pm)m}(\mathbf{x}) = 0,
\]

(3.3)

writing the spatial eigenfunctions \(\mathbf{h}_{(\pm)m}(\mathbf{x}) = (h_x(\pm)m(\mathbf{x}), h_y(\pm)m(\mathbf{x}))^T\), with the allowed wavenumber modulus \(\omega_m n/c\). Subtracting the product of the spatial differential equation (3.3), with the temporal eigenfunctions in each component, from the general differential equation (3.1), and factoring out the common spatial eigenfunctions leads to the corresponding temporal equation:

\[
\left(\frac{\partial^2}{\partial t^2} + \omega_m^2\right) \varphi_{(\pm)m}(t) = 0,
\]

(3.4)

where the two two-element temporal eigenfunctions are defined by \(\varphi_{(\pm)m}(t) = (\varphi_x(\pm)m(t), \varphi_y(\pm)m(t))^T\).

Our formulation of the problem is mathematically general, unlike more usual approaches [15, §7.1, pp. 269–273; 14, pp. 38–39]. It is more usual to think in terms of a "linearly polarized system" wherein only one light-wave element is selected from the outset. In such a case, it is only necessary to keep one spatial eigenfunction like \(h_x(\pm)_{m}(\mathbf{x})\) or \(h_y(\pm)_{m}(\mathbf{x})\) and keep only a scalar temporal wave. More generally, however, energy may be distributed between both spatial components, assuming that they have common eigenfrequencies and support standing waves of equal wavenumber. For simplicity, we will later be most interested in the case when the spatial wave elements are identical in each spatial component, i.e. \(h_x(\pm)_{m}(\mathbf{x}) = h_y(\pm)_{m}(\mathbf{x})\) for every eigenmode \(m\), and in the two spatial solutions \(\pm\). These configurations represent a nonbirefringent medium wherein the effective index of refraction \(n\) is the same for both electric-field components, or a multicavity arrangement with identical cavities such as we discuss below.

The second-order spatial differential equation (3.3) admits two wavenumber solutions in general, which we denote as \(\pm\). In rectangular geometries, the spatial eigenfunction is limited to just one allowed linear combination of
the two \( h_{j(\pm)m}(x) \) in each element \( j \). However, in this work it is necessary that the geometry and boundary conditions be sufficiently general to allow both solutions and linear combinations, giving independent \( \pm \) eigensolutions: \( h_{j(\pm)m}(x) \neq h_{j(-\pm)m}(x) \), in each eigenmode \( m \) and element \( j \). Independence between the \( \pm \) wavenumber eigenfunctions in every polarized element gives independence in all of the elements of the spatial eigenvector \( \mathbf{h}_{(\pm)m}(x) \). Also, since we are interested here in perturbations of the temporal eigenfunctions that arise from the nonlinear polarization term, which produce frequency splitting between the \( \pm \omega_m \) eigensolutions, we retain the full generality with two temporal solutions in each spatial element, and assume incoherence between the \( \exp(\pm i \omega_m t) \) components.

Although it turns out to be a purely academic exercise, it is interesting to understand the larger context of the temporal differential equation (3.1). Equation (3.1) is a general differential equation for waves not particular to light, but applicable to any two-element wave system. Two simple degenerate single-element waves with independent spatial components, as in a two-cavity arrangement, satisfy such a differential equation, and the idea generalizes simply to multiple-cavity arrangements. Spatial and temporal differential equations, like Equations (3.3) and (3.4), define the possible components of the spatial and temporal eigenfunctions, linear combinations of \( h_{j(\pm)m}(x) \) and \( \varphi_{j(\pm)m}(t) \) in each cavity, for each of the allowed eigenmodes in the two solutions denoted by \( \pm \). The temporal differential equation for the \( n_e \)-element resonator is written as in Equation (3.4), but with the two \( n_e \)-element solution vectors defined: \( \varphi_{(\pm)m}(t) = (\varphi_{1(\pm)m}(t), \varphi_{2(\pm)m}(t), \ldots, \varphi_{n_e(\pm)m}(t)) \).

A physical example of an \( n_e \)-element resonator can be constructed with \( n_e \) similar cavities for simple single-element waves with corresponding spatial positions in the cavities "wired" together, with unintrusive wires. Figure 1 sketches in outline a two-element resonator based upon two cavities and a three-cavity resonator. Degenerate standing-wave solutions between the elements lead to nonvanishing time-averaging pair products at the corresponding spatial positions. The wires allow for interference between the solution-vector elements at the corresponding spatial locations. The time-averaged pair products that result from the interference are assumed to determine the possible modifications in the medium at each spatial location for defining nonlinear feedback effects that modify the modes through the nonlinear polarization term. Degenerate solutions are most easily found if the cavities are identical, though this is not strictly necessary. With the introduction of a nonadjoint perturbation \( \mathbf{P}(x, t) \) there is a natural doubling of the number of elements as real and imaginary parts evolve independently. This work is mainly interested in nonrectangular geometries, as suggested by the 2D triangularly shaped cavities shown in the figure.
Fig. 1 Outline of two-cavity and three-cavity resonators. Spatially corresponding positions in the cavities of each resonator are "wired" together by unintrusive wires to allow interference between the wave elements to give intensities, like the light Stokes parameters, as functions of position to define nonlinear feedback processes in the resonator medium.
The second-order differential equation in \( n_e \) independent elements has two distinct eigenfrequencies \( \pm \omega_n \), corresponding to the two independent temporal waves. Writing Equation (3.4) as two coupled first-order differential \( n_e \)-element vector equations gives \( 2n_e \) equations written as one \( 2n_e \times 2n_e \) matrix equation:

\[
\left( i \mathbf{1} \cdot \frac{\partial}{\partial t} - \omega_m \mathbf{M}(\eta) \right) \varphi_m(t, \eta) = 0,
\]

where \( \mathbf{1} \) denotes the identity matrix of implied order, and the solution vector contains the two independent temporal eigenfunctions from the second-order differential equation (3.4) in its elements, i.e. \( \varphi_m(t, \eta) = (\varphi_{(+)m}(t)^T, \varphi_{(-)m}(t)^T)^T \), or \( \varphi_m(t, \eta) = (\varphi_{1(+)m}(t), \ldots, \varphi_{ne(+)m}(t), \varphi_{1(-)m}(t), \ldots, \varphi_{ne(-)m}(t))^T \). The driving matrix \( \mathbf{M}(\eta) \) is a constant \( 2n_e \times 2n_e \) matrix with two distinct eigenvalues \( \pm 1 \), each occurring \( n_e \) times. The matrix \( \mathbf{M}(\eta) \) can be written in terms of the \( 2n_e \times 2n_e \) matrix of eigenvectors \( \mathbf{U}(\eta) \) containing the \( 2n_e \) eigenvectors of \( \mathbf{M}(\eta) \) in its columns:

\[
\mathbf{M}(\eta) = \mathbf{U}(\eta) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{U}(\eta)^{-1}.
\]

As long as the eigenvalues of \( \mathbf{M}(\eta) \) are all \( \pm 1 \), the squared matrix is always the identity: \( \mathbf{M}(\eta) \cdot \mathbf{M}(\eta) = \mathbf{I} \), so applying the first-order time derivative twice gives the correct second-order equation, Equation (3.4). Since there must be equal numbers of \( \pm \) eigensolutions, \( \text{tr} \mathbf{M}(\eta) = 0 \), the trace being preserved in similarity transformations.

The spatial-temporal eigensolutions of Equation (3.5) can be written explicitly using the exponential of a matrix:

\[
\varphi_m(t, \eta) = \exp[-i \omega_m \mathbf{M}(\eta) t] \varphi_m(t_0),
\]

where \( \varphi_m(t_0) \) defines the initial conditions, or

\[
\tilde{\mathbf{E}}_m(x, t) = \exp[-i \omega_m \mathbf{M}(\eta) t] \tilde{\mathbf{h}}_m(x),
\]

where the \( 2n_e \)-element spatial-temporal solution vector is defined as the \( n_e \)-element generalization of the two component equation: \( \tilde{\mathbf{E}}_m(x, t) = (\tilde{\mathbf{E}}_{(+)m}(x, t)^T, \tilde{\mathbf{E}}_{(-)m}(x, t)^T)^T \) and

\[
\tilde{\mathbf{E}}_{(\pm) m}(x, t) = (\varphi_{1(\pm)m}(t, \eta) h_{1(\pm)m}(x), \varphi_{2(\pm)m}(t, \eta) h_{2(\pm)m}(x), \ldots, \\
\varphi_{ne(\pm)m}(t, \eta) h_{ne(\pm)m}(x))^T.
\]
and similarly for the spatial eigenfunctions:

\[
\mathbf{h}_m(x) = \left( \mathbf{h}_{(+)}^m(x)^T, \mathbf{h}_{(-)}^m(x)^T \right)^T
\]

\[
= (h_{1(+)}^m(x), \ldots, h_{n_{e(+)}}^m(x), h_{1(-)}^m(x), \ldots, h_{n_{e(-)}}^m(x))^T.
\]

The initial conditions \(\varphi_m(t_0)\) have been absorbed into the unnormalized spatial eigenfunctions:

\[
\tilde{\mathbf{h}}_m(x) = (\varphi_{1(\pm)}^m(t_0), h_{1(\pm)}^m(x), \varphi_{2(\pm)}^m(t_0), h_{2(\pm)}^m(x), \ldots,
\]

\[
\varphi_{n_{e}(\pm)}^m(t_0), h_{n_{e}(\pm)}^m(x))^T.
\]

which are also solutions to the linear spatial differential equation (3.3). The matrix \(\mathbf{M}(\eta)\) acts independently of the initial conditions in defining the temporal propagation in each vector element of \(\mathbf{E}_m(x, t)\) by forming linear combinations of products of the two independent and assumed incoherent temporal functions \(e^{\pm i\omega_m t}\) in the \(2n_e\) independent spatial eigenfunctions. The time dependence factors out of the matrix exponential on substituting the eigenvector matrix expansion for \(\mathbf{M}(\eta)\) from Equation (3.6), and simplifies to the convenient form

\[
\mathbf{E}_m(x, t) = \left( \frac{\exp(-i\omega_m t)}{2} [\mathbf{M}(\eta) + 1] - \frac{\exp(i\omega_m t)}{2} [\mathbf{M}(\eta) - 1] \right) \mathbf{h}_m(x)
\]

\[
= [(\cos \omega_m t) \mathbf{1} - i(\sin \omega_m t) \mathbf{M}(\eta)] \mathbf{h}_m(x).
\]

The driving matrix \(\mathbf{M}(\eta)\), the matrix of eigenvectors \(\mathbf{U}(\eta)\), and the solution vector for the eigenfunctions \(\varphi_m(t, \eta)\) are all written with an implied dependence on the parameter vector \(\eta\), because there exists a locus of temporal eigenfunctions continuous in a finite-element parameter space. It is most interesting to analyze the set of allowed temporal eigenfunctions in this form, because we find that only a subset of the parameters are affected in transformations of the eigenmode intensities. All complex matrices \(\mathbf{M}(\eta)\) having \(n_e\) eigenvalues of +1 and \(n_e\) eigenvalues of -1 are solutions of Equation (3.5). The \(2n_e\)-order matrix \(\mathbf{M}(\eta)\) is defined by \(8n_e^2\) arbitrary real parameters constrained by \(4n_e\) auxiliary real eigenvalue relations to give
4n_e(2n_e - 1) parameters. It is more meaningful to count the independent parameters defined by the Hermitian matrix \( M(\eta) \), as this form provides an orthogonal parameter basis. The general \( 2n_e \)-order Hermitian matrix \( M(\eta) \) is defined by \( 4n_e^2 \) arbitrary real independent parameters constrained the \( 2n_e \) auxiliary real eigenvalue relations, giving \( n_i = 2n_e(2n_e - 1) \) independent parameters: \( n_i = 2 \) for \( n_e = 1 \), \( n_i = 12 \) for \( n_e = 2 \), \( n_i = 30 \) for \( n_e = 3 \), etc. This estimate is also large, as it does not consider the effect of the repeated eigenvalues.

While trivial, the case of \( n_e = 1 \) is illustrative. For \( n_e = 1 \), the traceless Hermitian driving \( M(\eta) \) can be decomposed into the sum of the three mutually anticommuting \( 2 \times 2 \) Pauli spin matrices with the parameters in the three-element vector \( \eta \) as coefficients: \( M(\eta) = \eta_1 \sigma_1 + \eta_2 \sigma_2 + \eta_3 \sigma_3 \). For \( M(\eta) \) to be a valid driving matrix with \( M(\eta)^2 = 1 \) we have \( |\eta|^2 = \eta_1^2 + \eta_2^2 + \eta_3^2 = 1 \). Using the two free parameters in \( M(\eta) \), we form independent linear combinations of the two spatial eigenfunctions \( \hat{h}_{1(\pm)}(x) \) in the \( \sin \omega_n t \) temporal component in the two vector elements of \( \vec{E}_m(x, t) \) from Equation (3.9). The parameter basis acts independently of the initial conditions, giving independent spatial-temporal combinations in each of the vector elements of \( \vec{E}_m(x, t) \) as long as \( \hat{h}_{1(\pm)}(x) \not\propto \hat{h}_{1(\pm)}(x) \). The two independent parameters correspond to longitude and latitude coordinates on the unit square in the three-parameter space of \( \eta \).

For the case of \( n_e = 2 \), the four-element vector \( \vec{E}_m(x, t) \) represents the propagation of two independent monochromatic waves by the assumption of incoherence between the two temporal functions, \( \exp(\pm i \omega_n t) \). Taking \( U(\eta) = 1 \) or \( M(\eta) = \sigma_1 \otimes \sigma_0 \), we obtain the obvious solution in Equation (3.8) for \( \vec{E}_m(x, t) \): one polarization light state in the first two vector elements and another in the second two, with the Stokes intensities defined by the initial conditions. With a different choice for the eigenvectors in \( U(\eta) \), the two polarization states are viewed from a different basis, giving a distinguishable temporal variation from the equations (3.9), which consist of all possible linear combinations of the four independent spatial-temporal combinations between \( \sin \omega_n t \) and the elements of \( \hat{h}_m(x) \) added to \( (\cos \omega_n t)\hat{h}_m(x) \).

It is physically natural to expect a preference for the \( U(\eta) = 1 \) eigenmodes without birefringent effects or boundary conditions that couple the polarized components. A preference for separated temporal components seems natural in the multicavity designs from Figure 1. In the next section, we show that the change of basis represented by a different choice of eigenvectors produces a corresponding change in the measured Stokes intensities for the superposition of two monochromatic waves. While changing the basis may seem to be only a formalism, it does produce physically distinguishable effects with the introduction of the nonlinear-polarization term \( \mathbf{P}(x, t) \).
4. INTENSITY-DISTINGUISHABLE EIGENMODES

We proceed to isolate the subset of temporal wave solutions distinguishable in transformations of the coherency matrix of intensities, the intensity-distinguishable eigenmodes. The generalized intensities for the $n_e$-element resonator are defined by extension of the light Stokes parameters with the $n_e \times n_e$ coherency matrix

\[ J_m(x) = \left( [\tilde{\mathbf{E}}_{(+)}(x, t) + \tilde{\mathbf{E}}_{(-)}(x, t)] \otimes [\tilde{\mathbf{E}}_{(+)}(x, t) + \tilde{\mathbf{E}}_{(-)}(x, t)]^\dagger \right)_t \]

\[ = (1 1) \left( [\tilde{\mathbf{E}}_{m}(x, t) \otimes \tilde{\mathbf{E}}_{m}(x, t)]^\dagger \right)_t (1 1), \quad (4.1) \]

where the two vectors $\tilde{\mathbf{E}}_{(\pm)}(x, t)$ are the $n_e$-element solution vectors for the two spatial eigenmodes. The $2n_e$-element solution vector $\tilde{\mathbf{E}}_{m}(x, t)$ contains both the eigenmodes as defined earlier. The symbol 1 is used here to denote the $n_e \times n_e$ identity matrix in the $n_e \times 2n_e$ reducing matrix $(1, 1)$. We have summed the two eigensolutions assuming that the relevant processes do not distinguish between them, a usual assumption for light measurement processes as in Equation (1.1). The intensities are of course general functions of position $x$ in the domain of the resonator.

Substituting the eigenmode solution (3.9) into the coherency matrix (4.1) leaves only the dc terms in the expanded product under the time average, written one way:

\[ J_m(x) = \frac{1}{2}(1 1) \]

\[ \times \left( \mathbf{h}_m(x) \otimes \mathbf{h}_m(x)^\dagger + \mathbf{M}(\eta) [\mathbf{h}_m(x) \otimes \mathbf{h}_m(x)^\dagger] \mathbf{M}(\eta)^\dagger \right) \left( \begin{array}{c} 1 \\ 1 \end{array} \right). \]

\[ (4.2) \]

The coherency matrix $J_m(x)$, being Hermitian, is defined by $n^2_e$ real intensities. The four intensities for $n_e = 2$ correspond to the four Stokes parameters for light. The coherency matrix $J_m(x)$ is the sum of two Hermitian dyadics, outer products of a vector with its own transpose conjugate, corresponding to the two terms in Equation (4.2) after factoring the reducing matrices into the braces on the left and right. As a sum of two generally independent vector outer products, $J_m(x)$ is a Hermitian matrix with two generally nonzero eigenvalues. For $n_e = 2$, it is like the coherency matrix for partially polarized
light, which satisfies the familiar Stokes-parameter condition: \( s_0^2 \geq s_1^2 + s_2^2 + s_3^2 \). Partially polarized states arise because of the assumed independence of the two eigenmodes with eigenvalues \( \pm \omega_m \). Also, since \( J_m(x) \) of any order can always be written as the sum of two vector outer products, it can have no more that two nonzero eigenvalues. Thus for certain definitions of nonlinear polarizations that are dependent upon all of the intensities in \( J_m(x) \) in a higher-order multielement resonator \( (n_v > 2) \), \( n_v - 2 \) zero-eigenvalue constraints must be applied, effectively reducing the number of independent intensities.

The similarity transformation of \( J_m(x) \) with the unitary \( R(\xi) \), \( J'_m(x) = R(\xi)J_m(x)R(-\xi) \), maps the \( n_v^2 - 1 \) intensities [other than the trace of \( J_m(x) \)] into an equivalent set in a new basis. The unitary similarity transformation is factored into the larger matrix in curly braces in Equation (4.2) by any \( 2n_e \times 2n_e \) unitary matrix \( U(\xi) \) satisfying the commutation relation:

\[
R(\xi) \begin{pmatrix} 1 & 1 \\ \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} U(\xi).
\tag{4.3}
\]

With the similarity transformation of \( J_m(x) \), there is a corresponding change of basis in the solution basis; the spatial eigenfunction is modified: \( h'_m(x) = U(\xi)h_m(x) \), and the driving matrix transformed: \( M(\eta') = U(\xi)M(\eta)U(\xi)^\dagger \). The matrix \( M(\eta') \) defined by the new parameters \( \eta' \) is of course an allowed solution, since every pair of possible driving matrices is related by a similarity transformation.

As we saw from Equation (2.9), the commutation relation is satisfied by two unitary forms, (2.6) and (2.7), which generalize to arbitrary order \( n_v \). The general \( n_e \times n_e \) unitary transformation matrix \( R(\xi) \) is defined: \( R(\xi) = \exp[-iK(\xi)] \), based upon \( n_v^2 - 1 \) rotational parameters in the vector \( \xi \) using the \( n_v \times n_v \) traceless Hermitian matrix \( K(\xi) \). One general solution is evident from the fact that only the two Pauli spin matrices \( \sigma_0 \) and \( \sigma_2 \) commute with the reducing matrix. That solution is written \( U(\xi) = (\alpha \sigma_0 + \beta \sigma_2)U(0)^{(0)}(\xi) \), for coefficients \( \alpha \) and \( \beta \) with \( \alpha + \beta = 1 \). We take the case with \( \alpha = 1 \) and \( \beta = 0 \) as the representative unitary solution.

General higher-order solutions are obtained by expanding the exponential as a power series and realizing that successive \( n_e \) terms are independent by the Cayley-Hamilton theorem. The expansion into even and odd terms is represented by the matrix cosine and sine: \( \exp[-iK(\xi)] = \cos K(\xi) - i \sin K(\xi) \). Relating the two terms separately, we find the general solution: \( U(\xi) = (\alpha \sigma_0 + \beta \sigma_2) \otimes \cos K(\xi) - i(\alpha' \sigma_0 + \beta' \sigma_2) \otimes \sin K(\xi) \), where the complex coefficients satisfy: \( \alpha + \beta = 1 \) and \( \alpha' + \beta' = 1 \). However, in this general solution we realize only the one new unitary form, with \( \beta = \alpha' = 0 \); it is \( U(\xi) = U(1)^{(1)}(\xi) \).
Thus only a subset of the parameters $\eta$ in the driving matrix $M(\eta)$ and a corresponding subset of the wave solutions are modified by transformations of the intensities. We could use any of the unitary matrices $U(\xi)$ to generate that subset, but terms in $U^{(0)}(\xi)$ cancel due to the repeated eigenvalues in $M(\eta)$ to give a trivial result, so the next higher-order solution $U^{(1)}(\xi)$ is the lowest-order solution that can be used. We identify the subset of Hermitian driving matrices, $M(\xi)$ with the parameter vector $\zeta$, as the simplest complete set for $n_e > 1$ distinguishable in transformations of the intensities:

$$M(\xi) = U^{(1)}(\xi) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} U^{(1)}(\xi)^\dagger.$$  \hspace{1cm} (4.4)

The reduced parameter vector $\zeta$ consists of $n_2^2 - 1$ parameters, corresponding to the number of rotational degrees of freedom for transformations of the intensities. Substituting $U^{(1)}(\xi)$ from Equation (2.11) lets us write the arbitrary-order Hermitian driving matrix in terms of the general $n_e \times n_e$ traceless Hermitian matrix $K(\xi)$:

$$M(\xi) = \sigma_1 \exp[i2\sigma_2 \otimes K(\xi)] = \sigma_1 \otimes \cos 2K(\xi) - \sigma_3 \otimes \sin 2K(\xi).$$  \hspace{1cm} (4.5)

For $n_e = 2$, $U^{(1)}(\xi)$ is based upon the three parameters, and the resulting $M(\xi)$ is just $H_4(z)$ written in Equation (2.8).

Although the driving matrix $M(\xi)$ represents all of the intensities in the coherency matrix $f_{m}(\omega)$, the spatial variation of the intensities is affected by more than the reduced set of parameters $\zeta$. If the spatial eigenfunctions exhibit the same form in all their elements, i.e. $\hat{H}_{m(\pm)m}(\omega) \propto \hat{H}_{2(\pm)m}(\omega) \propto \cdots \propto \hat{H}_{n_r(\pm)m}(\omega)$ in each $\pm$ vector, then the spatial variation of the intensities factors out of the component matrices in the outer-product terms in the general coherency matrix. The two on-diagonal terms in $f_{m}(\omega)$, $\hat{H}_{(\pm)m}(\omega) \otimes \hat{H}_{(\pm)m}(\omega)^\dagger$, are approximately spatially constant, and the two off-diagonal terms, $\hat{H}_{(\pm)m}(\omega) \otimes \hat{H}_{\mp m}(\omega)^\dagger$, contain a scalar spatial function and its complex conjugate, respectively. The spatial constancy of the on-diagonal terms is strictly true for Fourier spatial eigenfunctions with $\hat{H}_{(\pm)m}(\omega) = \exp(\mp ik_m \cdot x)\hat{H}_{(\pm)m}(0)$, as we find with a rectangular geometry. Otherwise, the spatial variation of the terms is an approximation probably good to high order, as the wavelength of the modes is presumably very small compared to the size of the resonator.

It is straightforward to show that the Hermitian coherency matrix $f_{m}(\omega)$ can be written as a sum of three Hermitian coherency matrices multiplying
independent scalar spatial functions, an approximately constant spatial function multiplying the coherency matrix defined by the on-diagonal terms, and the real and imaginary parts of the common scalar spatial function multiplying coherency matrices in the off-diagonal terms. Each component coherency matrix can be determined by different intensities represented by three generally different parameter sets, within the constraints allowed by the total number of free parameters \( \eta \).

Our strict treatment of the two independent and incoherent temporal solutions \( \pm \omega_m \) in the coherency-matrix reduction (4.1)–(4.2), while rigorous for the complex solution, has an ambiguous meaning for the purely real-valued electric field. The reason is that the complex analytic extension by the Hilbert transform in time is not unique. The real parts of the two temporal solutions \( \exp(\pm i \omega_m t) \) both correspond to cosine waves, and therefore do not represent temporally incoherent signals. If the two solutions are not incoherent, it can be shown that the distinguishable parameter set \( \xi \) is reduced by one with an auxiliary relation. However, the derivation becomes relevant when a nonadjoint perturbation is added to the second-derivative operator of Equation (3.4). The added term can destroy this degeneracy, leading to a frequency splitting between the temporal components. We have therefore maintained the separation of the two temporal solutions and assumed incoherence between the temporal components, anticipating such a perturbation due to an added nonlinear polarization as we describe in Section 5.

5. QM ANALOG

The intensity-distinguishable set of standing light waves of a cavity or modes of a two-element resonator exhibits a multidimensional orthogonal parameter basis, like the four-dimensional Lorentz-transform-invariant space-time set. Introducing the \( n_e = 2 \) Hermitian solution (4.4) into the original differential equation (3.5), and writing it with the new variable \( p \), we find

\[
\begin{align*}
\left( i1 \frac{\partial}{\partial t} - \sum_{j=1}^{3} p_j \alpha_j - \mu_b \alpha_4 \right) \varphi_m(t, p) &= 0. \tag{5.1}
\end{align*}
\]

The new parameter vector \( p = (p_1, p_2, p_3)^T \) has units of frequency and is defined from the three polarization parameters \( \xi \):

\[
p_j = \xi_j \omega_m = \frac{\xi_{jm}}{\left(1 - |\xi_{m3}|^2\right)^{1/2}} \mu_b, \tag{5.2}
\]
where the new frequency is defined as

\[ \mu_b = \omega_m \left(1 - |\xi_m|^2 \right)^{1/2}. \] (5.3)

The parameter vector \( \xi_m \) is particular to the mode \( m \). By a particular mapping of the modal parameter-vector modulus \( |\xi_m|^2 \) with mode eigenfrequency \( \omega_m \), it is possible to obtain \( \mu_b \) constant for groups of eigenmodes, leading naturally to the idea that the new frequency \( \mu_b \) defines a resonator-state eigenfrequency. The new parameter vector obeys the identities

\[ |p/\mu_b|^2 = |\xi_m|^2/(1 - |\xi_m|^2) \] and \( (1 + |p/\mu_b|^2)^{1/2} = (1 - |\xi_m|^2)^{1/2}. \]

We have introduced the new index \( b \) to identify the eigenmode band to allow multiple spectral groupings of eigenmodes.

The differential equation for the set of intensity-distinguishable wave modes of a two-element resonator, Equation (5.1), has a basic resemblance to the Dirac equation for spin-\( \frac{1}{2} \) particle states from relativistic quantum mechanics. It exhibits the same degrees of freedom, but uses three polarization parameters rather than momenta as coefficients to the three matrices \( (\alpha_1, \alpha_2, \alpha_3) \), which are one of the forms for the Dirac alpha matrices, and with the resonator-state eigenfrequency substituting for the particle rest mass as the coefficient of \( \alpha_4 \), which is the same as the Dirac beta matrix. Strictly the Dirac alpha matrices that are most commonly used are a similarity transform of our three matrices \( (\alpha_1, \alpha_2, \alpha_3) \), corresponding to a different choice for the basis for the temporal eigenfunctions than the one we have used and not indicating a physical difference; specifically,

\[ \varphi^{(\text{Dirac})}_m(t, p) = \begin{pmatrix} 1 & 0 \\ 0 & i1 \end{pmatrix} \varphi_m(t, p). \] (5.4)

The usual units for the energy, momentum, and mass in the Dirac equation must be converted to units of frequency by factors of \( \hbar \) (Planck's constant over \( 2\pi \)) and \( c \) (the speed of light).

In a mathematical sense, the correspondence in matrix form between the differential equation (5.1) and the Dirac equation is not trivial. As many as twelve independent parameters may be needed to represent the general two-element Hermitian driving matrix, which are coefficients of generally nonsimilar matrix forms. But only four of the five independent \( 4 \times 4 \) anticommuting Hermitian spin matrices are used to represent the intensity-distinguishable two-element standing-wave modes as in the Dirac equation.

But what relationship do the degrees of freedom for transformations of the intensities represented by the polarization parameters in two-element
light standing-wave states have with the momentum coordinates for particle
spin-$\frac{1}{2}$ states in quantum mechanics? The two systems appear to be totally
different, considering that quantum mechanics is a representation for the
physics of our universe, albeit abstract and indirect. Other work has found
optical analogs with quantum mechanics. We remember Hamilton's classical
optical analog for Newtonian physics, the Hamilton-Jacobi equation, which
became the basis for the Schrödinger theory [13, Section 10–8]. Milony has
recently presented an analog for the 1D Schrödinger equation using a
particular optical arrangement [17]. All such optical analogs rely on the power
of the mathematical representation, as many of the basic features of quantum
mechanics are embodied in the form of the equations. It perhaps seems odd
to think that light, which we think of as consisting of spin-1 quantum
particles, could exhibit a spin-$\frac{1}{2}$ property, but it does not exhibit spin-$\frac{1}{2}$
features in real-space, only in the generated polarization-parameter space.

While it is beyond the scope of this paper to outline fully an optical
analog, it is useful to describe in a limited way the two ingredients that
are needed to build a viable optical analog. One is the generalization of
distinguishable parameters to a distinguishable-parameter space. A
distinguishable-parameter space can be defined with a superposition of
modes that densely spans a range in the parameter space $p$, where the mode
density is defined in a physically relevant sense. A desirable mapping between
the eigenmode frequency and its distinguishable parameters can be obtained
with an appropriate modal parameter-selection effect.

The second needed ingredient is an interaction effect that mimics the
photon exchange process between states in quantum mechanics. Such an
effect can be introduced naturally through the nonlinear polarization in
Equation (3.1). By mimicking photon exchange between states of the system
we should naturally obtain what appears to be a Coulomb field. From the
second-quantization theory, we know that the electromagnetic field can be
viewed as the result of the quantized exchange of energy between normal
particle states through photons. Essentially, the statistical probability of
photon exchange between similar eigenstates is mathematically equivalent to
a perturbation of the normal potential field, and can therefore be interpreted
in that way. This formulation leads to the well-known generation of a
Coulomb field for the probabilistic exchange of photons between eigenstates

With an appropriate mapping of the parameter-vector modulus as a
function of the mode eigenfrequency we can obtain constant $\mu_b$ in Equation
(5.3) in each frequency band $b$. This selection effect corresponds to a
changing basis for the definition of the intensity parameters as a function of
the mode of eigenfrequency. A changing choice of electromagnetic-vector
basis can be effected by varying the linear combination of components as a function of frequency in each wire in a multi-cavity resonator depicted in Figure 1. The choice of basis has nontrivial meaning in nonlinear processes, which are formed from multiple serial products of the electromagnetic components and may be defined by or include the coherency matrix (4.1). Mappings akin to this have been well studied in the design of birefringent filters, which introduce a desirable polarization encoding as a function of light color. The parameter selection defines a trajectory as a function of wavenumber, probably not unlike the unclosed light polarization transformation cycles, which are discussed for possible application in quantum experiments [4, 5].

With such a mapping it is possible to define a resonator state as a unique function of the distinguishable parameters \( p \). We can adopt a notation to represent the complex distribution of modes similar to what we have used to represent the temporal solution vector for a single mode, i.e. \( \varphi_m(t, p) \rightarrow \tilde{\Psi}(t, p) \) in Equation (5.1). This differential equation for the resonator state defines the solution vector \( \tilde{\Psi}(t, p) \) as a unique complex function of the distinguishable parameters \( p \).

Complex coefficients in the superposition must be encoded uniquely in the modal intensity distribution, which can be accomplished by varying the amplitudes and spatial offsets for the modes. We recognize that for nonrectangular cavity geometries both ± spatial eigenfunctions \( h_{(±)}m(x) \) and linear combinations, which define the spatial offset, can be allowed. Therefore we believe it to be possible to define the spatial offset for the mode independent of its parameters. Intensities remain well defined by the coherency matrix as long as the quasimonochromatic approximation is applicable, which is the case if the frequency band is limited to avoid the first harmonic overtone, or if the band is less in width that its base frequency. While description of nonlinear interaction effects in nonrectangular resonator geometries is beyond the scope of this paper, suffice it to say that subsets of closely interacting disjoint groupings of eigenmodes are naturally obtained from the nonorthogonality properties of the eigenmode intensity spatial functions, which are generated by the spatial eigenfunctions squared.

The Fourier transform of the solution vector, \( \tilde{\Psi}(t, p) \), as a function of \( p \), gives the vector \( \Psi(t, q) \), where \( q \) represents the Fourier-conjugate parameter domain. In the equation for \( \Psi(t, q) \), the Fourier transform of Equation (5.1), the parameters \( p \) are transformed into \( q \)-derivative operators. This property of the solution gives the familiar operator representation for the Hamiltonian in quantum mechanics and Hamilton's relations between conjugate canonical variables. The units for the parameters \( p \) are frequency, and those for the Fourier-conjugate parameter domain \( q \) are temporal. In the analog the unit of distance is defined from the temporal unit by the speed of light. The
uncertainty principle is the property of "optical" resolution which always exists between Fourier-conjugate quantities, as between the quantum momentum and spatial domains. Normal particle statistics is a property of the natural symmetries of the eigenstates of the basic relativistic equations like Equation (5.1).

Besides the dynamical theory represented by the Dirac equation, quantum mechanics includes what appears to be a fairly independent quantized statistical interaction theory, which describes the allowed jumps between the quantum eigenstates with the creation or collapse of intermediate photon wave functions. An idea akin to quantization is not foreign to nonlinear optics. Nonlinear effects in high-Q oscillators have been modeled assuming that the system is always in an eigenstate, but is allowed to jump instantaneously from one eigenstate combination to a different degenerate eigenstate combination [1, Section 1.5, pp. 36-50].

Resonant coupling through the second-order nonlinear susceptibility naturally produces traveling waves at the mode difference frequency, and driving of modes by low-frequency traveling waves naturally produces modes at the summed frequency. With a third-order nonlinear susceptibility, we can obtain a state self-intensity-dependent refractive index as a nonlinear polarization that has the form of a scalar multiplier of the eigenstate, similar to the form of the potential term in the Dirac driving equation. The terms that enter into the third-order nonlinear polarization term are the intensities from Equation (4.1). Coupling the state amplitude with its perturbed eigenfrequency precludes divided eigenstates, superpositions not being generally admissible as solutions because the potential term is nonlinear. Divided states are precluded too in photo-electron exchanges in quantum mechanics; e.g., an electron in a hydrogen atom is always found in a single quantum state, and never in a superposition that contains a combination of multiple allowed eigenfunctions.

With the change of amplitude of the eigenstates, the basic driving equation is perturbed. Presumably it is possible to obtain essentially instantaneous jumps from one allowed eigenstate with one eigenfrequency to another allowed eigenstate with a different eigenfrequency and a different self-intensity-dependent refractive index, with the creation or dissipation of a corresponding induced traveling wave with the appropriate frequency and total energy differences in the process. Eigenstate–travaling-wave transitions being discretized leads us necessarily to interpret the transition amplitude (the second-order nonlinear polarization) statistically as a probability for exchange, as in quantum mechanics. Conservation of energy in the transition necessarily conserves both the mode intensity and the frequency difference, as in the photon energy-frequency rule $E = h \nu$ for the frequency $\nu$ in cycles per second, with $h$ Planck's constant.
In a high-$Q$ oscillator, a normal-frequency resonator state can only be strongly influenced by resonant interaction with a traveling wave that exhibits a matching parameter distribution in the spatial domain of the resonator, or equivalently that exhibits a nearly corresponding distribution in the 3D distinguishable parameter space $p$ or in the Fourier-conjugate domain $q$. The interaction should be quite selective, with the induced traveling wave influencing only matching normal resonator states. An important feature of this form of exchange is that the traveling-wave time delay for interaction must be made to equal the $q$ separation between otherwise identical eigenstates, independent of the traveling-wave frequency. It turns out that this condition is obtained by constraining the nonlinear polarization or by affecting a selective coherent interaction such that the interaction only couples modes of parameter-vector amplitude $|p|_3$ to traveling waves of frequency $\omega$ with $\omega = |p|_3$.

A formulation based upon creation and dissipation of intermediate traveling waves in a two-element resonator parallels the photo-exchange process between quantum-mechanical particle states in real 3D space, as we illustrate in Figure 2. Presumably a normal-frequency resonator state, like $a$ (short-dashed lines), can jump to a lower-frequency allowed state, creating a traveling wave having the right difference frequency in the process, like $c$ (dotted lines). The traveling wave $c$ exhibits a matching distinguishable-parameter distribution in the spatial domain of the resonator initially, as indicated by dotted lines lying close to the short-dashed lines for $a$ in the upper panel of the figure. It therefore also exhibits a nearly matching $q$ distribution by the uniqueness of the modal expansion for resonator state $a$, and so appears to be coincident with $a$ in $q$ initially in the $(t, q)$ diagrams in the lower panels. The traveling wave presumably evolves differently from the normal resonator state, progressing away from $a$ both in its spatial distribution in the domain of the resonator and in $q$ space. Later, finding itself in near spatial coincidence in the spatial domain of the resonator with the normal resonator state $b$, it can resonantly drive state $b$ to a higher-frequency allowed eigenstate, being dissipated in the process. It is not unreasonable to expect that the resonator states should be localized in $q$, since the probability for interaction must decrease with $|q|_3$ for propagation of traveling waves away from a source state, like isotropically emitted photons in 3D space.

A fundamental feature of the two-element resonator equation is its Lorentz-transform invariance. This allows us to define a "local" time and coordinates for any particular frame of reference, as we denoted by superscripts $[a]$ or $[b]$ in the figure. Since the basic differential equation describing the resonator-state evolution retains its form in any Lorentz-transform basis, we cannot distinguish the properties of the system transformed to a scaled time and modified coordinates different from the real-time and normal-
FIG. 2. Traveling-wave progression between normal-frequency resonator states in spatial domain of resonator (top) and polarization parameter $q$ space (bottom). State $a$ with spatial distribution of intensities (short-dashed lines) in the resonator emits a traveling wave $c$ (dotted lines). Traveling wave $c$ evolves differently from $a$ in the resonator and in the $q$ space describing the separation between similar eigenstates. When it exhibits a matching distinguishable-parameter distribution with state $b$ (long-dashed lines) it is absorbed.

polarization basis of the resonator. This unique feature of the solutions means that the relevant time coordinates for the solutions may differ from the time in the resonator. A relevant time basis is one particular to the eigenstates, like the time delay for photon exchange between the proton and the coupled electron in a combined “hydrogen-atom.” Resonator states exhibiting a large
velocity in $q$ compared to stationary states in the special rest frame of reference with respect to the resonator (exhibiting proper times equal to the resonator time) naturally evolve more slowly in the resonator due to the relativistic invariance of the basic equations.

REFERENCES

20 H. Poincaré, La lumière, in Cours de Physique Mathématique; reprinted in Polarized Light (William Swindell, Ed.), Dowden, Hutchinson, and Ross, Stroudsburg, Penn., 1892, paper 11.


Received 8 December 1995, final manuscript accepted 16 September 1996