From: SYMMETRIES IN SCIENCE II
Edited by Bruno Gruber and Romuald Lenczewski
(Plenum Publishing Corporation, 1986)

LIGHT AND GROUP THEORY*

E. C. G. Sudarshan

Center for Particle Theory
University of Texas
Austin, Texas 78712

1. RADIATIVE TRANSFER AND ELECTRODYNAMICS

Light was recognized as the subtlest kind of matter quite a long time ago. Light has dominated human experience both in the natural form as well as when artifically produced. From the study of the propogation of light it was learned that light propogates rectilinearly.

When light is received from a great distance we find that it comes in the form of a beam but for artificially produced light we always see it as a pencil, originating at the source considered as a point: more generally, for an extended source we have a collection of pencils with vertices at the source. In both these cases it is as if the pencil (or beam) is composed of a multitude of individual rays each propogating rectilinearly. In practice, however, we never see individual rays; rather, we can think of arbitrarily narrow pencils which subtend smaller and smaller solid angles at the vertex.

The transmission of light energy by such complexes of pencils has been the subject of study in radiative transfer. In such studies we attempt to treat light as a collection of pencils, whose rays propogate rectilinearly. Since we can consider arbitrarily narrow pencils the radiant energy transfer can be reduced to integrodifferential equations.

We must expect substantial modification of the whole framework of radiative transfer when quantum ideas are invoked. Light has the attribute of polarization, which cannot be satisfactorily taken account of in conventional radiative transfer. Further since quantum theory of light demands that the wave aspect is essential, arbitrarily narrow pencils cannot be entertained; diffraction effects would then prevent rectilinear propogation.

In the context of quantum electrodynamics we have still further elaborations to be made. The field strengths now become quantum operators. However, that the first fundamental theorem of quantum optics asserts that every quantum illumination can be expressed as an ensemble of coherent wave fields with a possibly singular ensemble density. Each coherent field can

*Talk presented at the Symmetries in Science II Conference held at Southern Illinois University, Carbondale, March 23-26, 1986

for many purposes be thought of as a classical wavefield. So at least as far as lossless transmission in passive media is considered, the full quantum electrodynamic behaviour can be described as if we have an ensemble of classical wave fields. We shall rely on this fundamental theorem and treat optics as if it is classical wave optics with ensembles of wavefields as representative illuminations.

Light waves are transverse and therefore a full definition of the illumination involves the specification of polarization.

2. POLARIZED LIGHT AND TRANSFORMATIONS

When a parallel thin beam of light or a narrow paraxial pencil of light of a definite color is considered it has the additional quality of polarization. The polarization is transverse to the axis of propogation. We recognize two linearly independent states of polarization which may be chosen to be two linear orthogonal polarizations. The general state of polarization of a single wave would be a complex linear combination of these two. If the intensity of light is normalized the polarization state in general is a normalized complex combination; it therefore requires two complex numbers the sum of the absolute value being unity:

$$\underline{e} = \alpha_1 \underline{e}_1 + \alpha_2 \underline{e}_2 \; ; \quad |\alpha_1|^2 + |\alpha_2|^2 = 1.$$

Since measurable quantities are bilinear in the field, the overall phase of $\underline{e}$ is irrelevant (provided we have only one beam!) we may recognize that three real numbers the sum of whose squares is unity define the polarization state. For a pure state the polarization is thus specified not by a unimodular vector in a complex two-dimensional space but by a ray which is the equivalence class of all complex 2-vectors which differ from each other by a phase.

The special cases $\alpha_2/\alpha_1 = \pm i$ are circular polarizations. If α_2/α_1 is real we get a linear polarization. Both these are special cases of the general elliptic polarization.

In a general statistical light beam the polarization may not be complete and has to be represented by a polarization matrix ρ, which in turn is completely defined by the Stokes' parameters:

$$\underline{\Sigma} = \mathrm{tr}(\rho \underline{\sigma})$$

where σ_1, σ_2, σ_3 are 2x2 Pauli matrices. From the nonnegativity of ρ and its normalization

$$\mathrm{tr}(\rho) = 1 \; ; \quad \Sigma^2 \leq 1$$

with the equality being realized only for pure states with idempotent density matrices.

Consider the action of phase plates and optically active media on a beam of light and the manner in which the density matrix is transformed. To the extent that the propogation is lossless the set of transformations constitute the unitary unimodular group SU(2). Any such action can always be synthesized in terms of three preassigned kinds of objects using the familiar decomposition of a general rotation in 3-dimensions to be products of rotations respectively along z-axis, y-axis and z-axis. We may choose to synthesize any combination of elements affecting polarizations by three elements, respectively a phase plate, an optically active medium and a

phase plate:

$$A = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} e^{i\psi/2} & 0 \\ 0 & e^{-i\psi/2} \end{pmatrix}$$

In writing down this we have used a basis of orthogonal linear polarizations and absorbed overall phases in the representative of phase plates. The associated transformation of the density matrix is:

$$\rho \to A\rho\, A^{\dagger} = \rho' \quad ; \quad \mathrm{tr}(\rho') = \mathrm{tr}(\rho).$$

When absorption is involved like in a polarizer the matrix A is no longer unimodular:

$$\rho \to A\rho\, A^{\dagger} = \rho' \quad ; \quad \mathrm{tr}(\rho') < \mathrm{tr}(\rho).$$

More generally any linear map of $\rho \to \rho'$ which preserves the positivity is an acceptable map. If this is accomplished by beam splitting and recombination the map would belong to the subclass of "completely positive maps" with a standard form:

$$\rho \to \rho' = \sum_n A_n \rho A_n^{\dagger} \quad ; \quad \sum_n A_n^{\dagger} A_n - 1 \leq 0.$$

For the lossless case

$$\sum_n A_n^{\dagger} A_n = 1.$$

Given any two maps

$$\rho \to \sum_n A_n \rho A_n^{\dagger} \quad , \quad \rho \to \sum_m B_m \rho B_m^{\dagger}$$

we could form the convex combination

$$\rho \to \cos^2\alpha \Sigma A_n \rho A_n^{\dagger} + \sin^2\alpha \Sigma\, B_m\, \rho\, B_m^{\dagger}$$

and this could be rewritten in the form

$$\rho \to \sum_{\ell} C_{\ell}\, \rho\, C_{\ell}^{\dagger}$$

A map which cannot be expressed in this form is called extremal. The unitary maps

$$\rho \to A\rho A^{\dagger} \quad , \quad A^{\dagger}A = 1$$

are extremal. But we can also construct an extremal family of the form

$$\rho \to A_1 C\rho C A_1^{\dagger} + A_2 S\rho S A_2^{\dagger}$$

where A_1, A_2 are unitary and C, S are real matrices of the form

$$C = \begin{pmatrix} \cos\theta & 0 \\ 0 & 1 \end{pmatrix} \quad ; \quad S = \begin{pmatrix} \sin\theta & 0 \\ 0 & 0 \end{pmatrix} .$$

These two families exhaust the set of extremal completely positive maps. The antiunitary map

$$\rho \to A\rho^{T}A^{\dagger} \quad , \quad A^{\dagger}A = 1$$

is extremal but not completely positive and cannot be realized in terms of beam splitting and recombinations. The set of extremal not completely positive maps have been determined elsewhere.

3. RELATIVISTIC PARAXIAL PROPOGATION AND GROUPS

When we take account of relativity the light quantum appears as a realization of the Poincaré group with lightlike momenta. The irreducible realizations correspond to light quanta with positive or negative helicity. Thus circular polarizations are more natural in the context of relativistic transformations. Under any Lorentz transformation a left circularly polarized photon of a definite momentum goes into a photon of the same helicity and a new momentum with a phase factor which depends on the particular Lorentz transformation. A right circularly polarized photon would do the same but with the opposite phase.

If we temporarily neglect the polarization property and deal with scalar optics we could transform the scalar wave equation

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \right) \phi\,(xyz) = 0$$

into the form

$$\left(\frac{\partial^2}{\partial \tau \partial \sigma} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \right) \phi = 0$$

by introducing the light cone coordinates

$$\sigma = \frac{1}{2}\,(t-z) \quad , \quad \tau = \frac{1}{2}\,(t+z).$$

The elementary wave solution propogating along the z-axis has the alternate expressions

$$\exp\{i(k_3 z - \omega t)\} = \exp\{-i(\omega + k_3)\sigma - i(\omega - k_3)\tau\}$$

In the paraxial approximation $k_3 \approx \omega$ and for a pencil with small line width

we may replace $\omega + k$ by a quantity $2M$; when $\omega + k_3$ is constant we refer to it as being henochromatic, an idealization that we shall systematically use in paraxial optics. The coefficient of $-i\tau$ is

$$\omega - k_3 \approx \frac{k_1^2 + k_2^2}{2M} \quad ; \qquad k_1^2 + k_2^2 \ll M^2 \ .$$

Thus the differential equation for scalar wave propogation in the paraxial henochromatic idealization becomes

$$i\,\frac{\partial}{\partial \tau}\,\psi(x,y,\tau) \pm \frac{1}{2M}\left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2}\right)\psi(x,y,\tau)$$

reminiscent of the nonrelativistic Schrodinger equation with the henochromatic frequency M taking the place of the nonrelativistic mass.

The dynamical variables acting on the scalar wave amplitude are composed of $-i\frac{\partial}{\partial x}$, $-i\frac{\partial}{\partial y}$, x and y. The propogation "Hamiltonian" is quadratic in the momenta

$$\mathcal{H} = \frac{1}{2m}(p_1^2 + p_2^2);$$

$$p_1 = -i\frac{\partial}{\partial x} \quad , \quad p_2 = -i\frac{\partial}{\partial y} \quad .$$

We may consider the three bilinear quantities

$$\mathcal{H} = \frac{1}{2}(p_1^2 + p_2^2) \ ,$$

$$\mathcal{M} = (p_1 q_1 + q_2 p_2)$$

$$\mathcal{L} = \frac{1}{2}(q_1^2 + q_2^2)$$

which form the O(2,1) Lie algebra and the ten bilinear quantities

$$\mathcal{H} = \frac{1}{2}(p_1^2 + p_2^2), \ \mathcal{K}_1 = \frac{1}{2}(p_1^2 - p_2^2), \ K_2 = p_1 p_2$$

$$\mathcal{M} = (q_1 p_1 + p_2 q_2), \ \mathcal{N} = (q_1 p_1 - q_2 p_2)$$

$$\mathcal{J} = q_1 p_2 - q_2 p_1 \qquad \mathcal{N}_2 = (q_1 p_2 + p_2 q_1)$$

$$\mathcal{L} = \frac{1}{2}(q_1^2 + q_2^2) \qquad \mathcal{K}_1 = \frac{1}{2}(q^2 - q_2^2), \ \mathcal{K}_2 = q_1 q_2$$

constitute the O(3,2) Lie algebra. We shall see that the corresponding groups do play an important role in paraxial wave propogation in axial optical systems. In passing we also note that if the quantities q_1, q_2, p_1, p_2, 1 may be added to the Lie algebras to close them into their inhomegeneous semidirect product extensions. We shall see their relevance in the context of paraxial wave propogation in axial optical systems.

The inclusion of polarizations involve a nontrivial modification. If the scalar wave amplitude ϕ is replaced by a vector wave amplitude $\underline{V}$ (which may be taken to be the transverse divergence-free part of the vector potential) it is clear that multiplication by the transverse coordinates would take it outside the class of transverse vectors. Such an operator is therefore not defined but we must have a replacement for it so that we can deal with lens systems (see below). For this purpose one notes that the stabilizer group, that is the subgroup of the Lorentz group which leaves the light-like vector (1,0,0,1) is the Euclidean group E(2) made up of

$$G_1 = K_1 - J_2$$

$$G_2 = K_2 + J_1$$

$$J_3 \ .$$

In the scalar case

$$G_1 = (\omega + k_3)\, x_1 = 2M\, x_1$$

$$G_2 = (\omega + k_3)\, x_2 = 2M\, x_2$$

$$J_3 = 0.$$

Therefore

$$x_1 = \frac{1}{2M}\, G_1 \quad ; \quad x_2 = \frac{1}{2M}\, G_2.$$

But both in the spinless and in the spinning case G_1 and G_2 commute; we may therefore define as transverse coordinates

$$q_1 = \frac{1}{2M}\,(K_1 - J_2) = x_1 + \frac{1}{M}\, g_1 \;;$$

$$q_2 = \frac{1}{2M}\,(K_2 + J_1) = x_2 + \frac{1}{M}\, g_2 \;.$$

These augmentations of x_1 and x_2 preserve the transverse nature of the vector field $\underline{V}$ and contribute precisely the small longitudinal components:

$$g_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad ; \quad g_2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

As the frequency becomes higher the importance of these augmentations becomes less.

In either case the transverse coordinates are obtained from the generators of the stabilizer group.

4. GROUP THEORY OF AXIAL OPTICAL SYSTEMS

The passage of light through a system of optical elements like lenses, optically active media, or even free space displays realizations of the groups SO(2,1) and SO(3,2) and their inhomogeneous extensions. On encountering a thin spherical lens the geometrical thickness at any point (x_1, x_2) is of the form

$$\delta - \left(\frac{1}{2R_1} + \frac{1}{2R_2}\right)(x_1^2 + x_2^2)$$

and hence the wavefront suffers a phase shift

$$(n - 1)\delta - \frac{1}{2f}\,(x_1^2 + x_2^2)$$

where n is the refractive index of the lens medium and

$$f = (n - 1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right)$$

is the focal length of the lens. The constant term $(n - 1)\delta$ may be omitted without any loss of generality; for simplicity we have assumed that the lens is placed normal to the axis and centered. On the other hand the free particle propogation through a distance D is given by

$$D \quad (p_1^2 + p_2^2) \; / \; 2M.$$

These may therefore be identified with elements on the one-parameter subgroups generated by $\mathcal{L}$ and by $\mathcal{H}$. Since $\mathcal{L}$, $\mathcal{H}$ and $\mathcal{N}$ Lie algebra of SO(2.1) it follows that an exial system made up of a sucession of any number of spherical lenses separated b y arbitrary intervals can be identified with an element of the Lorentz group SO(2.1).

A more interesting question connected with the group properties of the axially symmetric system is that of synthesis: Given an arbitrary element of SO(2,1) how many lenses would be necessary and in what configuration to realize it? A more basic question is whether all such elements can be realized. These questions are nontrivial since while we can use both positive (convex) and negative (concave) lenses, the propogation distances D have to be chosen positive. By a straightforward but elaborate analysis we can show that all such SO(2,1) elements can be realized by a combination of at most three lenses with three separate propogation distances. In particular we could synthesize "propogation through a negative distance".

Noncentered axially symmetric lenses and lenses placed not precisely normal to the axis lead to terms linear in coordinates and momenta and therefore to the inhomogeneous extensions of the SO(2,1) group (which is better seen as the symplectic group Sp(2,R)).

Note that the linear terms in momenta and in coordinates are appropriate for a prism and for a phase plate.

If the lenses are not axially symmetric the optical phase shift on encountering a lens is a general quadratic form spanned by $\mathcal{L}$, $\mathcal{K}_1$ and $\mathcal{K}_2$. If the lenses are properly centered and oriented these suffice. Propogation introduces phases proportional to $\mathcal{H}$ and this together with $\mathcal{L}$, $\mathcal{K}_1$, $\mathcal{K}_2$ close on the Lie algebra SO(3,2). Any centered oriented system of lenses and free propogation realizes an element of SO(3,2). The synthesis question can be raised here also and can be answered in the affirmative.

The inhomogeneous versions of SO(3,2) or rather the symplectic Sp(4,R) correspond to excentric and mildly disoriented systems.

As a special case of such systems we may take cylindrical lenses and systems made by them to realize Sp(2,R) and its extension.

When the considerations change from scalar wave optics to vector wave optics we need to augment the coordinates by the matrices g_1 and g_2 so that all quantities are expressed in terms of q_1 and q_2.

5. LIGHT RAYS IN WAVE OPTICS

The fundamental representation of Sp(2,R) is by 2x2 matrices:

$$\mathcal{L} = \frac{1}{2}(q_1^2 + q_2^2) \quad \rightarrow \quad \begin{pmatrix} 0 & 0 \\ i & 0 \end{pmatrix}$$

$$\mathcal{H} = \frac{1}{2}(p_1^2 + p_2^2) \quad \rightarrow \quad \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}$$

$$\mathcal{N}_o = \frac{1}{2}(q_1 p_1 + p_2 q_2) \rightarrow \quad \begin{pmatrix} \frac{-i}{2} & 0 \\ 0 & +\frac{i}{2} \end{pmatrix}$$

These imaginary matrices and the corresponding realization of the group elements by real 2x2 matrices have an elementary geometrical interpretation in terms of light rays: q and p correspond to their spatial coordinates and their directions. A thin lens may be thought of as introducing an abrupt change in direction of the rays proportional to the distance of the point of incidence from the center; while free propogation maintains the direction but displaces the coordinate proportional to the slope! But how come we have rays coming up in wave optics?

Note that the symplectic groups Sp(2,R) and Sp(4,R) are there whether we think in terms of rays or not: the 2x2 and 4x4 matrices provide faithful realizations of Sp(2,R) and Sp(4,R) respectively.

We can obtain pencils of rays in wave optics if you realize that the two-point functions

$$\Gamma(x,y) = \langle \phi^{\dagger}(x)\, \phi(y) \rangle$$

furnish the mutual coherence functions in optics which propogate from one value of z to another: in quantum optics they are the normal ordered expectation values of field operators. Given $\Gamma(x,y)$ we can define the Wolf function by the Wigner-Moyal transformation:

$$W(q,p) = \int e^{-ip\xi}\, \Gamma(q + \tfrac{1}{2}\xi,\ q - \tfrac{1}{2}\xi)\, d\xi \ .$$

From the properties of $\Gamma(x,y)$ we can deduce that the Wolf function is always real but it is not always nonnegative. Any time W(q,p) becomes negative typical wavelike effects (diffraction, interference) can be expected.

Given any pencil of light with a Wolf function W(q,p) which may or may not be nonnegative, the effect of encountering an optical element like a lens or a segment of free propogation the new pencil of light is represented by a new Wolf function W'(q,p). The new pencil is obtained from the old pencil by pointwise ray transformation:

$$W'(q',p') = W(q,p).$$

We have thus the simple but startling result that the pencils in wave optics are transferred by an optical system in a manner simply described by geometrical optical transformation of the rays.

When mirrors are introduced we may generalize these considerations by path ordering the propogation and viewing the axis as being reflected by the mirrors, with a corresponding redefinition of the light cone coordinates. With this understanding a paraxial beam continues to be paraxial.

6. SQUEEZED LIGHT AND ESOTERICA

In the considerations so far the system elements were passive and lossless. The mutual coherence functions behave as convex cones, in the sense that whenever $\Gamma^{I}(x,y)$ and $\Gamma^{II}(x,y)$ are mutual coherence functions

$$\lambda_1 \Gamma^{I}(x,y) + \lambda_2 \Gamma^{II}(x,y) \ ; \ \lambda_1 > 0,\ \lambda_2 > 0 \ ,$$

is also an acceptable mutual coherence function. Identical convex cone property obtains for the Wolf functions. We may therefore deal with generating elements of these convex cones. The corresponding Wolf functions describe elementary pencils of light.

So far we have discussed the transformations of one such pencil into another by optical elements. These transform vector potentials into vector potentials and are therefore of the form of general linear canonical coordinate transformations. But from mechanics we know of more general canonical transformations. For a single mode we may work in terms of the creation and annihilation operators and ask for linear canonical transformations amongst them. We write

$$a \to \Lambda_{11}\, a + \Lambda_{12} a^{\dagger} \; ; \quad a^{\dagger} \to \Lambda_{21}\, a + \Lambda_{22}\, a^{\dagger} \; .$$

with

$$\Lambda_{11}^{*} = \Lambda_{22} \; , \; \Lambda_{12}^{*} = \Lambda_{21} \; ; \; \Lambda_{11}\Lambda_{22} - \Lambda_{21}\Lambda_{12} = 1 \; .$$

These are again Sp(2,R) transformations and can be generated by the matrices

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \; , \; \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \; , \; \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

This Lie algebra is realized by the generators

$$\frac{1}{4}\,(a^{\dagger}a + aa^{\dagger}) \; , \; \frac{i}{4}\,(a^{\dagger 2} - a^{2}) \; , \; \frac{1}{4}\,(a^{\dagger 2} + a^{2})$$

On the field ϕ (vector potential) and the conjugate momentum π (electric field) the transformations generated by SL(2,R) are real and include the natural time evolution mixing ϕ and π, as the scale transformations changing ϕ and π by reciprocal factors and the hyperbolic transformations between ϕ and π. Among these the scale transformations are most interesting since the uncertainty $\Delta\phi$ in the field variable ϕ, for example, could be made arbritarily small though $\Delta\phi.\Delta\pi$ could not be reduced. This process is called "squeezing". The simplest case of squeezing arises when we pass from a medium of high dielectric constant to one of low dielectric constant, though this is not easily amenable to experiment. In nonlinear optics such squeezed states have been identified and studied.

When mode mixing occurs we have more general linear canonical transformations amongst them. For two degrees of freedom we would have Sp(4,R), and, more generally for n degrees of freedom we would get Sp(2n,R). In addition to considering quantum electrodynamics of dielectrics which, in relation to empty space, realize squeezing transformations the optics of moving dielectrics realize one parameter subgroups of these more general transformations.

7. REMARKS

These considerations in optics are only some of the context in which group theory arises most naturally. There are many more which we have not discussed. One important example is the modification of a generalized Gaussian Schell beam: such a beam is defined by a mutual coherence function which is the exponential of a general quadratic form. Propogation changes this into another such beam; and these transformations have their own invariant, and the exponent transforms like a symmetric second rank tensor of the O(3,2) group. Similarly the propogation in an optical fibre with a refractive index varying quadratically with distance from center.

In addition to their mathematical elegance, group theoretic methods would also be useful in synthesis and design of systems.

ACKNOWLEDGEMENT

This work was supported by the Department of Energy gran number DE-FG05-85ER40200.