

Group theoretical methods in optics

N MUKUNDA

Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560012, India

Abstract. Scalar Fourier optics is concerned with the passage of paraxial light beams through ideal optical systems. It is well known that the action of the latter on the former can be given in the framework of the two- and four-dimensional real symplectic groups. It is shown here that, based on an analysis of the Poincaré symmetry of the complete Maxwell equations in the front form, a natural representation for paraxial Maxwell beams emerges, which moreover shows the way to a generalization of scalar to vector Fourier optics preserving the group structure of ideal optical systems. Properties of generalized rays, and the usefulness of some pseudo-orthogonal groups in the treatment of Gaussian Schell-model beams, are also brought out.

Keywords. Vector Fourier optics; first order optical systems.

I wish to describe here some work done recently by R Simon, E C G Sudarshan and myself on the application of group theoretical methods to a variety of problems in paraxial optics (Sudarshan *et al* 1983, Mukunda *et al* 1983). Starting from the Poincaré invariant Maxwell field equations of electromagnetism, one arrives at paraxial optics either in the ray or in the wave picture, making appropriate simplifications and approximations on the way (Bacry 1984). In the ray picture: we first suppress the vector nature of the field quantities, thus losing the polarization degree of freedom, and arrive at the Poincaré invariant scalar wave equation; for monochromatic waves with a fixed temporal frequency, this leads to the Helmholtz equation; in the eikonal approximation, where wave aspects are lost, one gets geometrical rays with simple propagation laws; and finally restricting these rays to form a narrow beam one arrives at the domain of paraxial ray optics. In the wave picture: one limits oneself to solutions of the scalar wave equation which are superpositions of plane waves all having a predominantly longitudinal wave-vector and very small transverse components, to reach the paraxial limit. Of course in this picture both polarization and simple rays are given up but typical wave aspects are retained.

In discussing the passage of paraxial scalar waves through so-called ideal optical systems, an elegant and practically useful group theoretical description has been found by Bacry and Cadilhac (1981). For monochromatic paraxial scalar waves, the wave equation becomes an equation determining evolution along the system axis—say the z -axis—of the distribution of field values over a transverse plane. When such a wave passes through a (linear) optical system, the system can be represented by a corresponding operator which acts on the incoming transverse field distribution to produce the outgoing one. Writing $\psi(x_{\perp})$ for the scalar field over a transverse plane, we can say:

$$\psi_{\text{output}}(x_{\perp}) = (\text{operator representing the system}) \psi_{\text{input}}(x_{\perp}) \quad (1)$$

One can now define an ideal optical system: the corresponding operator is the exponential of i times a hermitian and at most quadratic expression in the four hermitian operators x_{\perp} and $p_{\perp} = -i\nabla_{\perp}$. Examples are free propagation sections, prisms, thin lenses, and arbitrary combinations of these. Now x_{\perp} and p_{\perp} obey the canonical commutation relations.

$$[x_a, p_b] = i\delta_{ab}, [x_a, x_b] = [p_a, p_b] = 0; \quad a, b = 1, 2. \quad (2)$$

From this it follows that the operators representing ideal optical systems are elements of a group. In particular, leaving aside linear dependences on x_{\perp} and p_{\perp} in the exponent, one has the basic results: each axially symmetric system corresponds to an element of the group $SL(2, R) = Sp(2, R)$, while a system not necessarily possessing axial symmetry corresponds to an element of the group $Sp(4, R)$. The significance of these correspondences is that sequential action by several optical systems is represented by the product of the associated group elements in the same order.

These attractive results led us to the following natural questions:

(a) Can this group-theoretical description of optical systems be maintained consistently in the framework of the complete Maxwell equations, so that the polarization degree of freedom can be handled properly?

(b) Can the concept of rays of light be generalized so as to include the typical wave aspects of interference and diffraction; if so how do ideal optical systems act on such rays? How do group elements act on ray parameters?

The analysis leading to the answers to these questions can conveniently be presented in three parts:

(A) We start with the complete system of Maxwell equations and express its Poincaré invariance in the so-called front form. This is then exploited to develop a natural representation for the most general paraxial solution of Maxwell's equations. Thus we learn how to describe paraxial Maxwell beams with proper account of polarization.

(B) The rules of action of ideal optical systems on scalar waves—the domain of scalar Fourier optics—are then generalized so as to apply to vector waves, taking care to see that the association of systems with group elements is maintained (Mukunda *et al* 1985a).

(C) Within the framework of second order optical coherence theory we recall a recently proposed definition of generalized pencils of light rays (Sudarshan 1979a, b, 1981), capable of describing typical wave phenomena, and see how the group theoretic analysis extends to them.

To these ideas, we may append the following: for a particular class of light beams, the so-called Gaussian Schell model beams, one can develop an efficient and convenient description based on the pseudo-orthogonal Lorentz group $SO(2, 1)$ in the axially symmetric case and de Sitter group $SO(3, 2)$ in the general asymmetric case (Simon *et al* 1984, 1985).

We now describe each of the above briefly.

(A) The equations describing any relativistically invariant physical system maintain their forms under the action of the inhomogeneous Lorentz, or Poincaré, group \mathcal{P} . The effects of the elements of the group on the dynamical variables of the system are determined by ten infinitesimal generators, corresponding to the ten parameters of \mathcal{P} . The most familiar description of the generators is in the so-called instant form (Dirac 1949), wherein one uses a complete set of dynamical variables describing physical

conditions at all points of three-dimensional space at a common time. The generators of \mathcal{P} are then listed in the sequence: P^μ , generating space-time translations; \mathbf{J} , a three-vector generating spatial rotations; and \mathbf{K} , a three-vector generating pure Lorentz transformations. These ten generators have only a geometrical role or significance—they become dynamical quantities themselves only if one has a canonical formalism in classical theory, or one passes to quantum mechanics.

For problems of paraxial wave optics a more convenient way to display the action and generators of \mathcal{P} is the so-called front form (Dirac 1949). In place of the usual time and space coordinates $x^0 = ct$, \mathbf{x} we introduce the combinations.

$$\tau = \frac{1}{2}(x^0 + x^3), \sigma = x^0 - x^3, x_\perp. \quad (3)$$

A front is a hyperplane in space-time with a constant value of τ —we may call it “the front τ ”. Corresponding to the choice (3) of space-time coordinates, one rearranges the Poincaré generators as follows:

$$J_3; G_\perp = (K_1 - J_2)/2, (K_2 + J_1)/2; P_\perp; \\ M = (P^0 + P^3)/2; K_3 - \tau(P^0 - P^3); \quad (4a)$$

$$\mathcal{H} = P^0 - P^3; H_\perp = K_1 + J_2, K_2 - J_1. \quad (4b)$$

The seven generators listed in (4a) give rise to transformations of \mathcal{P} that map the front τ onto itself; the remaining ones in (4b) generate transformations that move the front. In particular the combination \mathcal{H} shifts a front while keeping it parallel to itself.

A particular feature of the rearrangement (4) of the Poincaré generators is that the six generators J_3, G_\perp, P_\perp, M drawn from (4a), along with \mathcal{H} taken from (4b), obey commutation relations corresponding to a two-space one-time Galilei algebra (Sudarshan and Mukunda 1974). In this algebra, \mathcal{H} plays the role of “energy”—it causes shifts in τ but not in σ or x_\perp —while M is like the “mass”. This Galilei structure is embedded in the overall Poincaré structure and is exposed by the front-form. Now the Galilei algebra with non-zero “mass” always contains within it, as is well known from nonrelativistic dynamics, quantities obeying the canonical or Heisenberg type commutation relations (Sudarshan and Mukunda 1974). These quantities are the spatial translation generators P_\perp and the boost generators G_\perp divided by the “mass” M . So if we define

$$Q_\perp = \frac{1}{M} G_\perp \quad (5)$$

and recall that M commutes with both G_\perp and P_\perp , we find as a consequence of the Poincaré commutation relations:

$$[Q_a, P_b] = i\delta_{ab}, [Q_a, Q_b] = [P_a, P_b] = 0. \quad (6)$$

Through this remark, we can trace the validity of the relations (2), which are in turn the basis for the group theoretic treatment of optical systems within the scalar theory, to the Poincaré group \mathcal{P} . Indeed in the case of the scalar wave equation one can easily obtain expressions for all the Poincaré generators in the front form, and one then discovers (Sudarshan *et al* 1983; Mukunda *et al* 1983):

$$\text{Scalar wave equation: } \begin{aligned} Q_\perp &= X_\perp, \\ P_\perp &= p_\perp = -i\nabla_\perp. \end{aligned} \quad (7)$$

Thus we have a clue to the generalisation to go from scalar to vector Fourier Optics: replace x_{\perp} of the scalar theory by Q_{\perp} appropriate to the Maxwell equations. Before taking this up, however, we mention that the front form analysis also tells us the most compact way in which paraxial solutions of the full system of Maxwell equations may be presented. When we compute Q_{\perp} for the electromagnetic field, it turns out to have the form.

$$\text{Maxwell's equations: } Q_{\perp} = x_{\perp} + \frac{1}{k} \tilde{G}_{\perp}. \quad (8)$$

Here, \tilde{G}_{\perp} are two matrices which are three-dimensional if the (free) Maxwell field is described by its electric vector, and six-dimensional if described by both electric and magnetic vectors. We use the former, more economical, description, and then find:

$$\tilde{G}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \tilde{G}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & i & 0 \end{pmatrix}. \quad (9)$$

The k in (8) is the wave number: $k = w/c$. Then the most general quasi-monochromatic paraxial solution to Maxwell's equation has an electric vector $\mathbf{E}(x, t)$ which may be displayed as (Mukunda *et al* 1985a, b)

$$\mathbf{E}(x, t) = \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} = \exp[(i/k)(\tilde{G}_a P_a)] E_T(x, t)$$

$$E_T(x, t) = \begin{pmatrix} E_1 \\ E_2 \\ 0 \end{pmatrix}. \quad (10)$$

Subject to the transverse components E_1 and E_2 being analytic signals with narrow angular spectra, there is full freedom in the choice of the column vector E_T . For any choice of E_T , application of the standard operator following it in (10) yields an allowed field \mathbf{E} . Thus in particular E_1 and E_2 may be chosen independently while E_3 is then determined. A similar description is possible using the magnetic field \mathbf{B} or the vector potential \mathbf{A} in a suitable gauge.

One must appreciate a subtle difference between the scalar wave equation and the Maxwell system at this point. The former wave equation $\square^2 \psi = 0$ reduces in the (quasi)-monochromatic and paraxial limit to an equation determining the evolution of $\psi(x_{\perp}, z)$ with respect to z . In particular, ψ may be chosen as we wish (subject to having a narrow angular spectrum) on an "initial" transverse plane, which will then "evolve" with respect to z in a definite way. The Maxwell system in the corresponding limiting situation turns out to contain both equations of constraint on the x_{\perp} dependences of \mathbf{E} and \mathbf{B} , and equations for the evolution of \mathbf{E} and \mathbf{B} with respect to z . The Poincaré symmetry expressed in the front form shows us how to disentangle these aspects properly, and leads to the representation (10) for the \mathbf{E} -field.

(B) Equipped with the above results, we can state the rule to go from scalar to vector Fourier optics. Referring to (1), we find: if in scalar optics a given (linear, ideal) optical system is represented by an operator $\omega(x_{\perp}, p_{\perp})$,

$$\psi_{\text{out}}(x_{\perp}) = \omega(x_{\perp}, p_{\perp}) \psi_{\text{in}}(x_{\perp}), \quad (11)$$

then the *same* system is represented in vector Fourier optics by a matrix function $\Omega(x_{\perp}, p_{\perp})$ of x_{\perp} and p_{\perp} obtained from $\omega(x_{\perp}, p_{\perp})$ by the replacement $x_{\perp} \rightarrow Q_{\perp}$:

$$\begin{aligned} \mathbf{E}_{\text{out}}(x_{\perp}) &= \Omega(x_{\perp}, p_{\perp}) \mathbf{E}_{\text{in}}(x_{\perp}), \\ \Omega(x_{\perp}, p_{\perp}) &= \omega(Q_{\perp}, p_{\perp}). \end{aligned} \quad (12)$$

Since Q_{\perp} and p_{\perp} obey the same algebra, (6), as do x_{\perp} and p_{\perp} , this prescription automatically guarantees that the association of linear ideal optical systems with elements of the groups $SL(2, R)$ or $Sp(4, R)$, depending on their symmetry, is maintained in vector Fourier optics. Thus, for the class of systems we are concerned with, we may say that

$$\omega'(x_{\perp}, p_{\perp}) \omega(x_{\perp}, p_{\perp}) = \omega''(x_{\perp}, p_{\perp}) \quad (13)$$

implies

$$\Omega'(x_{\perp}, p_{\perp}) \Omega(x_{\perp}, p_{\perp}) = \Omega''(x_{\perp}, p_{\perp}). \quad (14)$$

(More directly, this is so because Q_{\perp} is a similarity transform of x_{\perp}). The inclusion of the polarization degree of freedom in no way spoils the applicability of group theoretical methods in Fourier optics.

As an example we mention the case of a thin lens. Here, the formula of scalar optics,

$$\psi'(x_{\perp}) = \exp\left(\frac{-ik}{2f} x_{\perp}^2\right) \psi(x_{\perp}) \quad (15)$$

goes over into the formula

$$\begin{aligned} \mathbf{E}'(x_{\perp}) &= \exp(-ik/2f) Q_{\perp}^2 \mathbf{E}(x_{\perp}) \\ &= \exp(-ik/2f) x_{\perp}^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x/f & y/f & 1 \end{pmatrix} \mathbf{E}(x_{\perp}). \end{aligned} \quad (16)$$

Elsewhere we have given a proof of the above prescription and a systematic development of Fourier optics for the Maxwell field. (Mukunda *et al* 1985a)

From the group theoretical point of view the following question is interesting: suppose a given axially symmetric ideal system corresponds to an element $S \in SL(2, R)$. Then what, if any, system corresponds to the inverse element $S^{-1} \in SL(2, R)$? More generally, we can ask if indeed there is an optical system corresponding to each element of $SL(2, R)$, or whether only a part of the group "shows up" in this correspondence. This question is meaningful because the available simple "building blocks" out of which more complex systems may be built are: the thin lens with focal length f corresponding to the group element.

$$L(f) = \begin{pmatrix} 1 & 0 \\ -k/f & 1 \end{pmatrix}, \quad f \neq 0; \quad (17)$$

and free propagation through a distance D corresponding to

$$F(D) = \begin{pmatrix} 1 & D/k \\ 0 & 1 \end{pmatrix}, \quad D \geq 0. \quad (18)$$

It is the restriction $D \geq 0$ that makes the answer to the question not quite trivial. It can however be shown (Sudarshan *et al* 1985) that *every* element $S \in SL(2, R)$ can indeed be

realised as an axially symmetric ideal optical system, and in the construction one finds that one never needs more than three lenses plus three free propagation sections. In the nonaxially symmetric case, we have been able to show that every $S \in \text{Sp}(4, R)$ can be synthesized using a *finite* number of lenses and free sections.

(C) We have mentioned that each ideal axially symmetric optical system stands in one-to-one correspondence with an element of the group $\text{SL}(2, R)$ which can be pictured as a real 2×2 matrix with unit determinant. However when dealing with the wave amplitude and the changes it undergoes, there is no obvious quantity that is acted upon by these 2×2 matrices directly. Such an object however emerges when one deals with generalized rays and their distribution function. For simplicity we ignore polarization, though its inclusion is no problem. In statistical optics the state of a (scalar) light beam is described by the 2-point correlation function denoted schematically by

$$\Gamma(1; 2) = \langle \psi(1)\psi(2)^* \rangle. \quad (19)$$

Here the entries 1 and 2 denote two sets of space-time arguments and the angular brackets stand for an ensemble average. For a (quasi) monochromatic stationary field the time variables can be omitted and one deals with the spatial 2-point function $\Gamma(\mathbf{x}_1; \mathbf{x}_2)$. Sudarshan (1979a, b, 1981) has introduced the Wigner-Moyal transform of Γ , calling it the Wolf function, and proposed that it be interpreted as the density or distribution function of generalised rays of light corresponding to the given statistical state. The Wolf-function $W(\mathbf{x}, \mathbf{p})$ is the Fourier transform of $\Gamma(\mathbf{x} + 1/2\mathbf{x}'; \mathbf{x} - 1/2\mathbf{x}')$ with respect to the "difference" coordinate \mathbf{x}' , bringing in the new variable \mathbf{p} ; and represents the strength of rays of light at the point \mathbf{x} in the direction \mathbf{p} . It is real but can take both positive and negative values, corresponding to light and dark rays; this is necessary to allow for interference and diffraction. The generalized rays of light have very simple propagation laws but lose none of the typical wave aspects of light. In the paraxial situation (Simon 1983), W becomes a function of x_\perp and p_\perp only, and can be computed as the Wigner-Moyal transform of the 2-point function on a transverse plane. Like the field amplitude, the Wolf function also evolves from one transverse plane to another, and suffers changes as the beam passes through an optical system. In fact if we arrange the components of x_\perp and p_\perp as a column vector q , say, then we have the result

$$W_{\text{out}}(q) = W_{\text{in}}(S^{-1}q), \quad (20)$$

$$S \in \text{SL}(2, R) \text{ or } \text{Sp}(4, R).$$

As expressed graphically by Bacry (1984): In some sense, Geometrical optics is Fourier optics (up to a sign). We may amplify: the sense is that of generalized rays.

Gaussian Schell model (GSM) beams

An unexpectedly rich area of application of group theoretical techniques in optics has turned out to be the study of a special class of fields known as the Gaussian Schell model fields. These are defined by the particular form the 2-point function (over a transverse plane) takes, and are of considerable practical interest. A beam is said to be of Schell model type if the normalised degree of coherence γ is translation invariant in a transverse plane. We have a GSM beam if furthermore the intensity distribution I , the normalised degree of coherence γ and the phase ϕ are all Gaussian, i.e., exponentials of quadratic forms. It is then easy to see that by the Wigner-Moyal transform we are led to

a Wolf-function which is likewise Gaussian: for GSM beams,

$$\begin{aligned} W(x_{\perp}, p_{\perp}) &= \exp[-(\text{positive definite quadratic form in } x_{\perp} \text{ and } p_{\perp})] \\ &= \exp[-q^T G q], \end{aligned} \quad (21)$$

where the 4×4 real symmetric matrix G completely characterizes the beam (Simon *et al* 1984, 1985). For axially symmetric fields, G is essentially a 2×2 matrix, while in general it is four dimensional. The characteristic features of G such as positive definiteness and some more subtle properties have been completely analysed. They lead to the following geometrical constructions: axially symmetric GSM fields can be represented in a one-to-one way as time-like vectors in a fictitious $2 + 1$ dimensional Minkowski space; and the action on them by symmetric ideal optical systems, symbolized by

$$G \rightarrow G' = S G S^T, \quad S \in \text{SL}(2, R), \quad (22)$$

is given by Lorentz transformations on these vectors. If one gives up the axial symmetry, then each field G can be represented in a fictitious $3 + 2$ dimensional de Sitter space as a certain second rank antisymmetric tensor, while the action on it by a system corresponding to some $S \in \text{Sp}(4, R)$ amounts to a de Sitter transformation.

These geometrical representations of GSM fields are practically useful because they allow a systematic classification of the fields and evaluation of the invariants associated with them. Many questions of physical interest, which are extremely cumbersome to treat in the traditional wave-optical language, become quite easy to answer in the combined generalized-ray and group theoretical language.

It is perhaps surprising, but yet gratifying, that we have been able to find so many uses for group theoretical methods in such a well-established subject as optics!

References

- Bacry H 1984 Group theory and paraxial optics. Invited talk at 13th Int. Colloq. on group theoretical methods, University of Maryland
- Bacry H and Cahillac M 1981 *Phys. Rev.* **A23** 2533
- Dirac P A M 1949 *Rev. Mod. Phys.* **21** 392
- Mukunda N, Simon R and Sudarshan E C G 1985a *J. Opt. Soc. Am.* **A2** 416
- Mukunda N, Simon R and Sudarshan E C G 1985b *J. Opt. Soc. Am.* **A** (in press)
- Simon R 1983 *Pramana* **20** 105
- Simon R, Sudarshan E C G and Mukunda N 1984 *Phys. Rev.* **A29** 3273
- Simon R, Sudarshan E C G and Mukunda N 1985 *Phys. Rev.* **A** (in press)
- Sudarshan E C G 1979a *Phys. Lett.* **A73** 269
- Sudarshan E C G 1979b *Physica* **A96** 315
- Sudarshan E C G 1981 *Phys. Rev.* **A23** 2802
- Sudarshan E C G and Mukunda N 1974 *Classical dynamics: A modern perspective* (New York: John Wiley) ch. 9
- Sudarshan E C G, Mukunda N and Simon R 1985 *Opt. Acta* (in press)
- Sudarshan E C G, Simon R and Mukunda N 1983 *Phys. Rev.* **A28** 2921