Optik 125 (2014) 968–978

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Aberrations in Maxwell optics

Sameen Ahmed Khan

Engineering Department, Salalah College of Technology (SCOT), Post Box No. 608, 211 Salalah, Oman

ARTICLE INFO

Article history:
Received 18 March 2013
Accepted 20 July 2013

PACS:
02
42.15.−i
42.90.+m

Keywords:
Vector wave optics
Foldy-Wouthuysen transformation
Wavelength-dependent effects
Beam propagation
Hamiltonian description

ABSTRACT

An exact treatment of beam optics, starting *ab initio* from the Maxwell's equations is presented. The starting point of this approach is a matrix representation of the Maxwell's equation in a medium with varying permittivity and permeability. Formal expressions are obtained for the paraxial and leading order aberrating Hamiltonians, without making any assumptions on the form of the varying refractive index. We derive the wavelength-dependent contributions at each order, starting with the lowest-order paraxial Hamiltonian. To illustrate the general theory, we consider the computations of the transfer maps for an axially symmetric graded-index medium. For this system, in the traditional approaches, one gets only six aberrations. In our formalism, we get all the nine aberrations permitted by the axial symmetry. The six aberrations coefficients of the traditional approaches get modified by the wavelength-dependent contributions and the remaining three are pure wavelength-dependent. It is very interesting to note that apart from the wavelength-dependent modifications of the aberrations, this approach also gives rise to the *image rotation*. The present study is the generalization of the traditional and non-traditional prescription of Helmholtz optics. In the low wavelength limit our formalism reproduces the Lie algebraic formalism of optics. The present study further strengthens the close analogies between the various prescriptions of light optics and charged-particle optics. The new formalism presented here, provides a natural framework to study beam-optics and polarization in a unified manner.

1. Introduction

Several formalisms are available for the scalar wave theory of optics. Oldest among these is based on the beam-optical Hamiltonian derived using the Fermat's principle. It is well-known that the Maxwell's theory of electromagnetism is the correct theory of light. So, it is natural to develop a beam-optics formalism based on the Maxwell's equations. This is often done using the Helmholtz equation. In this approach, one takes the *square-root* of the Helmholtz operator followed by an expansion of the radical [1–3]. It should be noted that the square-root approach reduces the original *boundary value problem* to a *first-order initial value problem*. This reduction is of great practical value, since it leads to the powerful systems or the Fourier optic approach [4]. However, the beam optical Hamiltonian in the square-root approach is *no* different from the geometric approach of the Fermat's principle. Moreover, the reduction process itself can never be rigorous or exact! Hence, there is a room for alternative procedures for the reduction and several reduction schemes are discussed in [5,6]. Of course any such reduction scheme is bound to lack in rigour to some extent, and the ultimate justification lies only in the results it leads to. The issue of the square-root can be elegantly circumvented by exploiting the algebraic similarities between the Helmholtz equation and the Klein–Gordon equation. Consequently, the Helmholtz equation is linearized using the Feshbach–Villars procedure originally developed for the Klein–Gordon equation [7]. This casts the Helmholtz equation to a Dirac-like form enabling the use of the Foldy–Wouthuysen expansion used in the Dirac electron theory [8]. This approach provides an alternative to the square-root approach. This formalism gives rise to wavelength-dependent contributions modifying the paraxial behaviour [5] and the aberration coefficients [6]. This is the *non-traditional* prescription of scalar wave optics. In the low wavelength limit it reproduces the traditional prescriptions based on the square-root and the Fermat's principle. The suggestion to employ the Foldy–Wouthuysen transformation technique in the case of the Helmholtz equation was mentioned in the literature as a remark [9,10]. It was only in the recent works, that this idea was exploited to analyze the quasiparaxial approximations for specific beam-optical systems [5,6]. The Foldy–Wouthuysen technique is ideally suited for the Lie algebraic approach to optics. With all these plus points, the powerful and ambiguity-free expansion, the Foldy–Wouthuysen transformation is still little used in optics [11,12]. The technique of the Foldy–Wouthuysen transformation results in what we call as the non-traditional prescriptions of

E-mail address: rohelakhan@yahoo.com
URLs: http://SameenAhmedKhan.webs.com/.

0030-4026/$ – see front matter © 2013 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.ijleo.2013.07.097

Helmholtz optics [6]. The algebraic machinery of this formalism is adopted from the machinery used in the *quantum theory of charged-particle beam optics*, based on the Dirac equation [13–21] and the Klein–Gordon equation [10,21]. A treatment of beam optics taking into account the anomalous magnetic moment is available in [22–28]. A diffractive quantum limit has been also considered for particle beams [29]. The emerging field of QABP, the *quantum aspects of beam physics*, has been recognized by a series of meetings by the same name [30–32].

The Helmholtz equation is derived from the Maxwell's equations. But it is approximate as one neglects the spatial and temporal derivatives of the permittivity and permeability of the medium. So, *no* prescription based on the Helmholtz equation can be exact! One has to take into account the vectorial nature of the light field and more specifically the fact that the Maxwell's is a *constrained* system of equations. So, it is very natural to look for a prescription fully based on the Maxwell's equations. The Maxwell's equations are coupled and it is difficult to deal with them in their usual form. The logical starting point is to cast the Maxwell's equations in a matrix form: a single entity containing all the four Maxwell's equations. The beam-optical Hamiltonian derived from the matrix representation of the Maxwell's equations has an algebraic structure very similar to the Dirac equation. This enables us to use the Foldy–Wouthuysen transformation technique. The formalism thus developed provides a deeper understanding of beam-optics, particularly in the wavelength-dependent regime. Moreover, it provides a framework to study beam-optics and light polarization in a unified manner.

From the very beginning, let us be very clear that the Dirac field and the electromagnetic field are two entirely different entities. But the resemblance in the underlying algebraic structure of the Dirac equation and the matrix representation of the Maxwell's equation can be exploited to carry out some relevant calculations leading to well-established results. The Foldy–Wouthuysen transformation was historically developed for the Dirac equation, particularly for understanding its nonrelativistic limit [33–42]. The technique initially developed for the spin-1/2 particles was extended to the spin-0 and the spin-1 particles [43], and even generalized to the case of arbitrary spins [44]. The Foldy–Wouthuysen iterative diagonalization technique can be applied to certain types of equations, which have a particular algebraic structure. The problems under study need not be even quantum mechanical! Consequently, the Foldy–Wouthuysen technique has found to be applicable to a variety of problems, such as atomic systems [45,46]; synchrotron radiation [47] and derivation of the Bloch equation for polarized beams [48]. In the context of acoustics, comprehensive and mathematically rigorous accounts can be found in [49–54]. It has found applications in ocean acoustics as well [55]. A comprehensive account of the use of Foldy–Wouthuysen transformation in optics is available in [11,12]. At no stage, we are invoking any equivalence between the four-dimensional Dirac field and the six-dimensional electromagnetic field governed by the Maxwell's equations. We have used the Foldy–Wouthuysen transformation technique from a calculational purpose. In this paper, we develop the beam-optical formalism starting with the exact matrix representation of the Maxwell's equations in an inhomogeneous medium, taking into account the spatial and temporal variations of the permittivity and permeability. The required matrix representation and how it differs from the other representations is described in Section 2. Section 3 has the beam-optical formalism. Section 4 has the applications. Section 5 has our concluding remarks.

2. Matrix representation of Maxwell's equations

There are different matrix representations of Maxwell's equations which were derived with a different motivation [56–61]. Moreover, some of these are in vacuum or make use of a pair of matrix equations. In our formalism, we require a single matrix equation containing all the four Maxwell's equations, taking into account the spatial and temporal variations of the permittivity $\epsilon(\boldsymbol{r}, t)$ and permeability $\mu(\boldsymbol{r}, t)$. Such a representation was specifically developed for the beam-optical formalism [62]. The Maxwell's equations [63,64] in an inhomogeneous medium with sources are

$$
\begin{aligned}
&\nabla \cdot \boldsymbol{D}(\boldsymbol{r}, t) = \rho, \\
&\nabla \times \boldsymbol{H}(\boldsymbol{r}, t) - \frac{\partial}{\partial t}\boldsymbol{D}(\boldsymbol{r}, t) = \boldsymbol{J}, \\
&\nabla \times \boldsymbol{E}(\boldsymbol{r}, t) + \frac{\partial}{\partial t}\boldsymbol{B}(\boldsymbol{r}, t) = 0, \\
&\nabla \cdot \boldsymbol{B}(\boldsymbol{r}, t) = 0.
\end{aligned} \tag{1}
$$

The media is assumed to be linear throughout this study. That is $\boldsymbol{D} = \epsilon\, \boldsymbol{E}$, and $\boldsymbol{B} = \mu\, \boldsymbol{H}$, where $\epsilon = \epsilon(\boldsymbol{r}, t)$ and $\mu = \mu(\boldsymbol{r}, t)$ are the permittivity and the permeability of the medium. In general ϵ and μ vary with space and time. The speed of light in the medium is $v(\boldsymbol{r}, t) = 1/\sqrt{\epsilon(\boldsymbol{r}, t)\mu(\boldsymbol{r}, t)}$ and the refractive index of the medium is by $n(\boldsymbol{r}, t) = c/v(\boldsymbol{r}, t) = c\sqrt{\epsilon(\boldsymbol{r}, t)\mu(\boldsymbol{r}, t)}$. The resistance of the medium is given by $h(\boldsymbol{r}, t) = \sqrt{\mu(\boldsymbol{r}, t)/\epsilon(\boldsymbol{r}, t)}$. As we shall shortly see, it is advantageous to use the two derived laboratory functions $v(\boldsymbol{r}, t)$ and $h(\boldsymbol{r}, t)$ instead of $\epsilon(\boldsymbol{r}, t)$ and $\mu(\boldsymbol{r}, t)$. In terms of these functions, $\epsilon = 1/hv = n/ch$ and $\mu = h/v = hn/c$. Following the notation in [59–61], we use the Riemann–Silberstein vector [65,66] given by

$$
\boldsymbol{F}^{\pm}(\boldsymbol{r}, t) = \frac{1}{\sqrt{2}}\left(\sqrt{\epsilon(\boldsymbol{r}, t)}\boldsymbol{E}(\boldsymbol{r}, t) \pm \mathrm{i}\frac{1}{\sqrt{\mu(\boldsymbol{r}, t)}}\boldsymbol{B}(\boldsymbol{r}, t)\right). \tag{2}
$$

We further define,

$$
\Psi^{\pm}(\boldsymbol{r}, t) = \begin{bmatrix} -F_x^{\pm} \pm \mathrm{i}F_y^{\pm} \\ F_z^{\pm} \\ F_z^{\pm} \\ F_x^{\pm} \pm \mathrm{i}F_y^{\pm} \end{bmatrix}, \tag{3}
$$