International Journal of Theoretical Physics, Vol. 44, No. 1, January 2005 (© 2005)
DOI: 10.1007/s10773-005-1488-0

Wavelength-Dependent Modifications in Helmholtz Optics

Sameen Ahmed Khan[1]

Received September 29, 2004; accepted November 24, 2004

The Helmholtz wave equation is linearized using the Feshbach–Villars procedure used for linearizing the Klein–Gordon equation, based on the close algebraic analogy between the Helmholtz equation and the Klein–Gordon equation for a spin-0 particle. The Foldy–Wouthuysen iterative diagonalization technique is then applied to the linearized Helmholtz equation to obtain a Hamiltonian description for a system with varying refractive index. The Hamiltonian has a wavelength-dependent part absent in the traditional descriptions. Besides reproducing all the traditional quasi-paraxial terms, our method leads to additional contributions dependent on the wavelength. Applied to the axially symmetric graded-index fiber, this method results in wavelength-dependent modifications of the paraxial behavior and the aberration coefficients to all orders. Explicit expression for the modified aberration coefficients to the third order are presented. Sixth- and eighth-order Hamiltonians are also presented.

KEY WORDS: Scalar wave optics; Helmholtz equation; wavelength-dependent effects; beam propagation; Hamiltonian description; aberrations; graded-index fiber; mathematical methods of optics; Feshbach–Villars linearization; Foldy–Wouthuysen transformation.
PACS: 02,42.90.+m

1. INTRODUCTION

Historically, the scalar wave theory of optics (including aberrations to all orders) is based on Fermat's *principle of least time*. In this approach, the beam-optical Hamiltonian is derived using Fermat's principle. This approach is purely geometrical and works adequately in the scalar regime. Later on it was realized that the whole of optics is governed by Maxwell's equations. All the laws of geometrical optics can be deduced from Maxwell's equations (Born and Wolf, 1999). This deduction is traditionally done using the Helmholtz equation, which is derived from Maxwell's equations. In this approach, one takes the *square-root* of the Helmholtz operator followed by an expansion of the radical (Dragt *et al*., 1986;

[1] Middle East College of Information Technology (MECIT), Technowledge Corridor, Knowledge Oasis Muscat, Post Box 79, Al Rusayl 124, Muscat, Sultanate of Oman; e-mail: sakhan@mecit.edu.om.

Dragt, 1998). 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 system or the Fourier optic approach (Goodman, 1996). However, the beam-optical Hamiltonian in the square-root approach is *no* different from the geometrical approach of Fermat's principle. Moreover, the reduction process itself can never be claimed to be rigorous or exact.

The purpose of this report is to present an alternate procedure for the reduction, based on the close algebraic analogy between the Helmholtz equation and the Klein–Gordon equation for a spin-0 particle. Our approach, which uses the algebraic machinery of quantum mechanics, provides a 'natural' procedure for the reduction. Furthermore, our procedure gives rise to some interesting extra contributions, modifying the beam-optical Hamiltonian of geometrical optics. This results in the corrections to the beam optics even at the 'paraxial-level' (Khan *et al.*, 2002).

The Helmholtz equation governing scalar optics is algebraically very similar to the Klein–Gordon equation for a spin-0 particle. Exploiting this similarity, the Helmholtz equation is linearized in a procedure very similar to the one due to Feshbach–Villars for linearizing the Klein–Gordon equation. This brings the Helmholtz equation to a Dirac-like form enabling the procedure of the Foldy–Wouthuysen expansion used in the Dirac electron theory. This approach, which uses the algebraic machinery of quantum mechanics, was developed recently (Khan *et al.*, 2002), providing an alternative to the traditional *square-root* procedure. The formalism presented here gives rise to wavelength-dependent contributions modifying the aberration coefficients. The algebraic machinery of this formalism is very similar to the one used in the *quantum theory of charged-particle beam optics*, based on the Dirac (Jagannathan *et al.*, 1989; Jagannathan, 1990; Jagannathan, 1993) and the Klein–Gordon (Khan and Jagannathan, 1995) equations respectively. The detailed account for both of these is available in (Jagannathan and Khan, 1996). A treatment of beam optics taking into account the anomalous magnetic moment is available in (Conte *et al.*, 1996; Jagannathan and Khan, 1997; Khan, 1997; Jagannathan and Khan, 1998). A complete coverage to the new field of *Quantum Aspects of Beam Physics* (QABP), can be found in the proceedings of the series of meetings under the same name (Chen, 1999).

Using our approach, we derive general expressions for the Hamiltonians without assuming any specific form for the refractive index. These Hamiltonians are shown to contain extra wavelength-dependent contributions, which arise very naturally in our approach. We apply the general formalism to the specific examples:

A. *Medium with constant refractive index*. This example is essentially for illustrating some of the details of the machinery used.
B. *Axially symmetric graded-index medium*. This example is used to demonstrate the power of the formalism. The traditional approaches

give six third-order aberrations. Our formalism modifies these six aberration coefficients by wavelength-dependent contributions. Sixth- and eighth-order Hamiltonians are also derived for this system. All the associated machinery used in this formalism is described in the text and the appendices.

The traditional beam optics (in particular, the Lie algebraic formalism of light beam optics) (Dragt *et al.*, 1986; Dragt, 1998) is completely obtained from our approach in the limit wavelength, $\bar{\lambda} \longrightarrow 0$, which we call the *traditional limit* of our formalism. This is analogous to the *classical limit* obtained by taking the reduced Planck's constant $\hbar = h/2\pi \longrightarrow 0$, in the quantum prescriptions. The scheme of using the Foldy–Wouthuysen machinery in this formalism is very similar to the one used in the *quantum theory of charged-particle beam optics*, developed in recent years (Jagannathan *et al.*, 1989; Jagannathan, 1990, 1993; Khan and Jagannathan, 1995; Jagannathan and Khan, 1996, 1997, 1998; Conte *et al.*, 1996; Khan, 1997). There too one recovers the classical prescriptions (in particular, the Lie algebraic formalism of charged-particle beam optics (Turchetti *et al.*, 1989; Todesco, 1999)) in the limit $\bar{\lambda}_0 \longrightarrow 0$, where $\bar{\lambda}_0 = \hbar/p_0$ is the reduced de Broglie wavelength and p_0 is the design momentum of the system under study.

In this report we focus on the Hamiltonian description of beam optics, as is customary in the traditional prescriptions of beam optics. This also enables us to relate our formalism with the traditional prescriptions, such as the Lie algebraic formalism and the quantum-like approach (Fedele and Man'ko, 1999).

2. TRADITIONAL PRESCRIPTIONS

In the traditional scalar wave theory for treating monochromatic quasiparaxial light beam propagating along the positive z-axis, the z-evolution of the optical wave function $\psi(\mathbf{r})$ is taken to obey the Schrödinger-like equation

$$i\bar{\lambda}\frac{\partial}{\partial z}\psi(\mathbf{r}) = \widehat{H}\psi(\mathbf{r}), \tag{1}$$

where the optical Hamiltonian $\widehat{H}$ is formally given by the radical

$$\widehat{H} = -\left(n^2(\mathbf{r}) - \widehat{\mathbf{p}}_{\perp}^2\right)^{1/2}, \tag{2}$$

and $n(\mathbf{r}) = n(x, y, z)$ is the varying refractive index. In beam optics, the rays are assumed to propagate almost parallel to the optic-axis, chosen to be z-axis, here. That is, $|\widehat{\mathbf{p}}_{\perp}| \ll p_z \approx 1$ and $|n(\mathbf{r}) - n_0| \ll n_0$. The refractive index is the order of unity. Let us further assume that the refractive index varies smoothly around the constant background value n_0 without any abrupt jumps or discontinuities. For a medium with uniform refractive index, $n(\mathbf{r}) = n_0$ and the Taylor expansion of the radical is

$$
\begin{aligned}
(n^2(\mathbf{r}) - \widehat{\mathbf{p}}_\perp^2)^{1/2} &= n_0 \left\{1 - \frac{1}{n_0^2}\widehat{\mathbf{p}}_\perp^2\right\}^{1/2} \\
&= n_0 \Bigg\{1 - \frac{1}{2n_0^2}\widehat{\mathbf{p}}_\perp^2 - \frac{1}{8n_0^4}\widehat{\mathbf{p}}_\perp^4 - \frac{1}{16n_0^6}\widehat{\mathbf{p}}_\perp^6 \\
&\qquad - \frac{5}{128n_0^8}\widehat{\mathbf{p}}_\perp^8 - \frac{7}{256n_0^{10}}\widehat{\mathbf{p}}_\perp^{10} - \cdots\Bigg\}. \qquad (3)
\end{aligned}
$$

In the earlier expansion, one retains terms to any desired degree of accuracy in powers of $(\frac{1}{n_0^2}\widehat{\mathbf{p}}_\perp^2)$. In general, the refractive index is not a constant and varies. The variation of the refractive index, $n(\mathbf{r})$ is expressed as a Taylor expansion in the spatial variables x, y with z-dependent coefficients. To get the beam-optical Hamiltonian, one makes the expansion of the radical as before, and retains terms to the desired order of accuracy in $(\frac{1}{n_0^2}\widehat{\mathbf{p}}_\perp^2)$ along with all the other terms (coming from the expansion of the refractive index $n(\mathbf{r})$) in the phase-space components up to the same order. In this expansion procedure, the problem is partitioned into paraxial behavior + aberrations, order-by-order.

3. THE FOLDY–WOUTHUYSEN FORMALISM

In the traditional scheme, the purpose of expanding the beam-optical Hamiltonian $\widehat{H} = -(n^2(\mathbf{r}) - \widehat{\mathbf{p}}_\perp^2)^{1/2}$ in a series, is to understand the propagation of the quasi-paraxial beam in terms of a series of approximations (paraxial + nonparaxial). In relativistic quantum mechanics too, one has the problem of understanding the behavior in terms of nonrelativistic limit + relativistic corrections terms in the quasi-relativistic regime, order-by-order. For the Dirac equation (which is first order in time) this is done using the Foldy–Wouthuysen transformation leading to an iterative diagonalization technique (Foldy and Wouthuysen, 1950; Bjorken and Drell, 1964). For the Klein–Gordon equation (which is second order in time) this is done using the same Foldy–Wouthuysen technique after linearizing it with respect to time, and thus bringing it to a Dirac-like form, following the Feshbach–Villars method (Feshbach and Villars, 1958). Here, we follow a procedure very similar to the one used for linearizing the Klein–Gordon equation *via* the Feshbach–Villars linearizing procedure (Feshbach and Villars, 1958). The resulting Feshbach–Villars-like form has an algebraic structure very similar to the Dirac equation. This enables us to make an expansion using the Foldy–Wouthuysen transformation technique well known in the Dirac electron theory (Foldy and Wouthuysen, 1950; Bjorken and Drell, 1964). The resulting expansion reproduces the aforementioned expression in (3) as it should. Furthermore, it gives rise to a set of wavelength-dependent contributions. The analogy between the Helmholtz wave equation and the Klein–Gordon equation suggests naturally a similar technique for treating the scalar wave theory of light beams. Though the suggestion to employ the Foldy–Wouthuysen technique in the case of the Helmholtz equation existed in the literature as a remark (Fishman and McCoy, 1984), it has been exploited only now, to analyze the quasi-paraxial approximations (Khan *et al.*, 2002). The formalism presented here is an elaboration of the work initiated in (Khan *et al.*, 2002), providing an alternative to the traditional square-root approach.

We start with the wave equation in the rectilinear coordinate system.

$$\left\{\boldsymbol{\nabla}^2 - \frac{n^2(\mathbf{r})}{v^2}\frac{\partial^2}{\partial t^2}\right\}\Psi = 0. \tag{4}$$

Let

$$\Psi = \psi(\mathbf{r})e^{-i\omega t}, \quad \omega > 0, \tag{5}$$

then

$$\left\{\boldsymbol{\nabla}^2 + \frac{n^2(\mathbf{r})}{v^2}\omega^2\right\}\psi(\mathbf{r}) = 0. \tag{6}$$

At this stage, we introduce the wavization,

$$-i\lambdabar\boldsymbol{\nabla}_\perp \longrightarrow \widehat{\mathbf{p}}_\perp, \qquad -i\lambdabar\frac{\partial}{\partial z} \longrightarrow p_z, \tag{7}$$

where $\lambdabar = \lambda/2\pi$ is the reduced wavelength, $c = \lambdabar\omega$ and $n(\mathbf{r}) = c/v(\mathbf{r})$. It is to be noted that $pq - qp = -i\lambdabar$. This is similar to the commutation relation, $(pq - qp) = -i\hbar$, in quantum mechanics. In our formalism, $\lambdabar$ plays the same role that is played by the reduced Planck's constant, $\hbar$ in quantum mechanics. The traditional beam-optics formalism is completely obtained from our formalism in the limit $\lambdabar \longrightarrow 0$. Then, we get,

$$\left\{\left(-i\lambdabar\frac{\partial}{\partial z}\right)^2 + (\widehat{\mathbf{p}}_\perp^2 - n^2(\mathbf{r}))\right\} \quad \psi(\mathbf{r}) = 0. \tag{8}$$

Next, we linearize Eq. (8) following a procedure similar to the one which gives the Feshbach–Villars (Feshbach and Villars, 1958) form of the Klein–Gordon equation. To this end, let

$$\begin{pmatrix}\psi_1(\mathbf{r})\\ \psi_2(\mathbf{r})\end{pmatrix} = \begin{pmatrix}\psi(\mathbf{r})\\ -i\frac{\lambdabar}{n_0}\frac{\partial}{\partial z}\psi(\mathbf{r})\end{pmatrix}. \tag{9}$$

Written as a first-order system, the Helmholtz equation now reads

$$-i\frac{\bar{\lambda}}{n_0}\frac{\partial}{\partial z}\begin{pmatrix}\psi_1(\mathbf{r})\\ \psi_2(\mathbf{r})\end{pmatrix} = \begin{pmatrix}0 & 1\\ \frac{1}{n_0^2}(n^2(\mathbf{r})-\widehat{\mathbf{p}}_\perp^2) & 0\end{pmatrix}\begin{pmatrix}\psi_1(\mathbf{r})\\ \psi_2(\mathbf{r})\end{pmatrix}. \tag{10}$$

Next, we make the transformation,

$$\begin{aligned}\begin{pmatrix}\psi_1(\mathbf{r})\\ \psi_2(\mathbf{r})\end{pmatrix} \longrightarrow \Psi^{(1)} = \begin{pmatrix}\psi_+(\mathbf{r})\\ \psi_-(\mathbf{r})\end{pmatrix} &= M\begin{pmatrix}\psi_1(\mathbf{r})\\ \psi_2(\mathbf{r})\end{pmatrix}\\ &= \frac{1}{\sqrt{2}}\begin{pmatrix}\psi_1(\mathbf{r})+\psi_2(\mathbf{r})\\ \psi_1(\mathbf{r})-\psi_2(\mathbf{r})\end{pmatrix}\\ &= \frac{1}{\sqrt{2}}\begin{pmatrix}\psi(\mathbf{r})-i\frac{\bar{\lambda}}{n_0}\frac{\partial}{\partial z}\psi(\mathbf{r})\\ \psi(\mathbf{r})+i\frac{\bar{\lambda}}{n_0}\frac{\partial}{\partial z}\psi(\mathbf{r})\end{pmatrix}\end{aligned} \tag{11}$$

where

$$M = M^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \qquad \det M = -1. \tag{12}$$

It is to be noted that the transformation matrix M is independent of z. For a monochromatic quasi-paraxial beam (in forward direction), with leading z-dependence $\psi(\mathbf{r}) \sim \exp\{in(\mathbf{r})z/\bar{\lambda}\}$. Then

$$\begin{aligned}\psi_+ &\sim \frac{1}{\sqrt{2}}\left\{1+\frac{n(\mathbf{r})}{n_0}\right\}\psi(\mathbf{r}),\\ \psi_- &\sim \frac{1}{\sqrt{2}}\left\{1-\frac{n(\mathbf{r})}{n_0}\right\}\psi(\mathbf{r}).\end{aligned} \tag{13}$$

Since, $|n(\mathbf{r})-n_0| \ll n_0$, we have $\psi_+ \gg \psi_-$. Consequently, Eq. (8) can be written as

$$\begin{aligned}i\bar{\lambda}\frac{\partial}{\partial z}\begin{pmatrix}\psi_+(\mathbf{r})\\ \psi_-(\mathbf{r})\end{pmatrix} &= \widehat{\mathrm{H}}\begin{pmatrix}\psi_+(\mathbf{r})\\ \psi_-(\mathbf{r})\end{pmatrix},\\ \widehat{\mathrm{H}} &= -n_0\sigma_z + \widehat{\mathcal{E}} + \widehat{\mathcal{O}}\\ \widehat{\mathcal{E}} &= \frac{1}{2n_0}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\}\sigma_z\\ \widehat{\mathcal{O}} &= \frac{1}{2n_0}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\}(i\sigma_y),\end{aligned} \tag{14}$$

where σ_y and σ_z are, respectively, the y and z components of the triplet of Pauli matrices,

$$\boldsymbol{\sigma} = \left[\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right]. \tag{15}$$

It is to be noted that the even-part and odd-part in Hamiltonian (14) differ only by a Pauli matrix. This simplifies the computations a lot as we shall see, shortly. The details of the Feshbach–Villars linearizing procedure for the Klein–Gordon equation are available in Appendix A.

The square of the Hamiltonian in (14) is

$$\widehat{\mathrm{H}}^2 = (n^2(\mathbf{r}) - \widehat{\mathbf{p}}_\perp^2), \tag{16}$$

as expected. Thus, we have taken the square root in a different way. Our procedure of taking the square root is based on the experience of the Klein–Gordon equation and has certain advantages over the traditional procedure of directly taking the square root.

The purpose of casting Eq. (8) in the form of Eq. (14) will be obvious now, when we compare the latter with the Dirac equation

$$\begin{aligned} i\hbar \frac{\partial}{\partial t}\begin{pmatrix} \Psi_{\mathrm{u}} \\ \Psi_{\mathrm{l}} \end{pmatrix} &= \widehat{H}_{\mathrm{D}} \begin{pmatrix} \Psi_{\mathrm{u}} \\ \Psi_{\mathrm{l}} \end{pmatrix} \\ \widehat{H}_{\mathrm{D}} &= m_0 c^2 \beta + \widehat{\mathcal{E}}_{\mathrm{D}} + \widehat{\mathcal{O}}_{\mathrm{D}} \\ \widehat{\mathcal{E}}_{\mathrm{D}} &= q\phi \\ \widehat{\mathcal{O}}_{\mathrm{D}} &= c\boldsymbol{\alpha} \cdot \widehat{\boldsymbol{\pi}}, \end{aligned} \tag{17}$$

where 'u' and 'l' stand for the upper and lower components, respectively, and

$$\boldsymbol{\alpha} = \begin{bmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{bmatrix}, \qquad \beta = \begin{bmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{bmatrix}, \qquad \mathbb{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{18}$$

To proceed further, we note the striking similarities between Eq. (14) and Eq. (17). In the nonrelativistic positive energy case, the upper components Ψ_{u} are large compared to the lower components Ψ_{l}. The odd ($\widehat{\mathcal{O}}$) part of ($\widehat{H}_{\mathrm{D}} - m_0 c^2 \beta$), anticommuting with β couples the large Ψ_{u} to Ψ_{l}, while the even ($\widehat{\mathcal{E}}$) part commuting with β, does not couple them. Using this fact, the well known Foldy–Wouthuysen formalism of the Dirac electron theory (see, *e.g.*, Bjorken and Drell, 1964) employs a series of transformations on Eq. (17) to reach a representation in which the Hamiltonian is a sum of the nonrelativistic part and a series of relativistic correction terms; $|c\widehat{\boldsymbol{\pi}}|/m_0 c^2$ serves as the expansion parameter and the nonrelativistic part corresponds to an approximation of order up to $|c\widehat{\boldsymbol{\pi}}|/m_0 c^2$. The terms of higher order in $|c\widehat{\boldsymbol{\pi}}|/m_0 c^2$ constitute the relativistic corrections. Examining

Eq. (14), we conclude $\psi_+ \gg \psi_-$, and the odd operator $\widehat{\mathcal{O}}$, anticommuting with σ_z, couples the large ψ_+ with the small ψ_-, while the even operator $\widehat{\mathcal{E}}$ does not make such a coupling. This spontaneously suggests that a Foldy–Wouthuysen-*like* technique can be used to transform Eq. (14) into a representation in which the corresponding beam-optical Hamiltonian is a series with expansion parameter $|\widehat{\mathbf{p}}_\perp|/n_0$. The correspondence between the beam-optical Hamiltonian (14) and the Dirac electron theory is summarized in a table in Appendix B. The Foldy–Wouthuysen technique iteratively takes the field to a new representation, where the forward-propagating components get progressively *decoupled* from the backward-propagating components.

Application of the Foldy–Wouthuysen-*like* technique to Eq. (14) involves a series of transformations on it and after the required number of transformations, depending on the degree of accuracy, Eq. (14) is transformed into a form in which the residual odd part can be neglected and hence the upper and lower components (ψ_+ and ψ_-) are effectively decoupled. In this representation, the larger component (ψ_+) corresponds to the beam moving in the $+z$-direction and the smaller component (ψ_-) corresponds to the backward-moving component of the beam.

Using the correspondence between Eqs. (14) and (17) the Foldy–Wouthuysen expansion given formally in terms of $\widehat{\mathcal{E}}$ and $\widehat{\mathcal{O}}$ leads to the Hamiltonian

$$i\lambda\!\!\!^{-} \frac{\partial}{\partial z} |\psi\rangle = \widehat{\mathcal{H}}^{(2)} |\psi\rangle,$$

$$\widehat{\mathcal{H}}^{(2)} = -n_0\sigma_z + \widehat{\mathcal{E}} - \frac{1}{2n_0}\sigma_z \widehat{\mathcal{O}}^2. \tag{19}$$

To simplify the formal Hamiltonian we use, $\widehat{\mathcal{O}}^2 = -\frac{1}{4n_0^2}\{\widehat{\mathbf{p}}_\perp^2 + (n_0^2 - n^2(\mathbf{r}))\}^2$ and recall that $\widehat{\mathcal{E}} = \frac{1}{2n_0}\{\widehat{\mathbf{p}}_\perp^2 + (n_0^2 - n^2(\mathbf{r}))\}\sigma_z$. We are primarily interested in the forward-propagating beam; so we drop σ_z. Then the formal Hamiltonian in (19) is expressed in terms of the phase-space variables as:

$$\widehat{\mathcal{H}}^{(2)} = -n_0 + \frac{1}{2n_0}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\} + \frac{1}{8n_0^3}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\}^2. \tag{20}$$

The Foldy–Wouthuysen iterative procedure is described in Appendix B. The lowest-order Hamiltonian obtained in this procedure agrees with the traditional approaches, as it should.

To go beyond the expansions in (20), one goes a step further in the Foldy–Wouthuysen iterative procedure. To next-to-leading order the Hamiltonian is formally given by

$$i\lambda\!\!\!^{-} \frac{\partial}{\partial z} |\psi\rangle = \widehat{\mathcal{H}}^{(4)} |\psi\rangle,$$

$$
\begin{aligned}
\widehat{\mathcal{H}}^{(4)} &= -n_0\sigma_z + \widehat{\mathcal{E}} - \frac{1}{2n_0}\sigma_z\widehat{\mathcal{O}}^2 \\
&\quad - \frac{1}{8n_0^2}\left[\widehat{\mathcal{O}}, \left([\widehat{\mathcal{O}}, \widehat{\mathcal{E}}] + i\lambda\!\!\!^{-}\frac{\partial}{\partial z}\widehat{\mathcal{O}}\right)\right] \\
&\quad + \frac{1}{8n_0^3}\sigma_z\left\{\widehat{\mathcal{O}}^4 + \left([\widehat{\mathcal{O}}, \widehat{\mathcal{E}}] + i\lambda\!\!\!^{-}\frac{\partial}{\partial z}\widehat{\mathcal{O}}\right)^2\right\}.
\end{aligned}
\tag{21}
$$

As before, we drop the σ_z and the resulting Hamiltonian in the phase-space variable is

$$
\begin{aligned}
\widehat{\mathcal{H}}^{(4)} &= -n_0 + \frac{1}{2n_0}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\} \\
&\quad + \frac{1}{8n_0^3}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\}^2 \\
&\quad - \frac{i\lambda\!\!\!^{-}}{32n_0^4}\left[\widehat{\mathbf{p}}_\perp^2, \frac{\partial}{\partial z}(n^2(\mathbf{r}))\right] \\
&\quad + \frac{\lambda\!\!\!^{-2}}{32n_0^5}\left(\frac{\partial}{\partial z}(n^2(\mathbf{r}))\right)^2 \\
&\quad + \frac{1}{16n_0^5}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\}^3 \\
&\quad + \frac{5}{128n_0^7}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\}^4 .
\end{aligned}
\tag{22}
$$

The Hamiltonian thus derived has all the terms, which one gets in the traditional square-root approach. In addition, we also get the wavelength-dependent contributions. Before proceeding further, let us examine the leading order modifications to the paraxial Hamiltonian

$$
\widehat{\mathcal{H}}^{(p)} = -n_0 + \frac{1}{2n_0}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n^2(\mathbf{r})\right)\right\} - \frac{i\lambda\!\!\!^{-}}{32n_0^4}\left[\widehat{\mathbf{p}}_\perp^2, \frac{\partial}{\partial z}(n^2(\mathbf{r}))\right]. \tag{23}
$$

It is clear that the paraxial Hamiltonian has an extra term, the commutator term. This term is always present as long as the refractive index has inhomogeneities. Such a term does not arise in the traditional prescriptions.

The details of the various transforms and the beam-optical formalism being discussed here turns out to be a simplified analog of the more general formalism developed recently for the *quantum theory of charged-particle beam optics*, (Jagannathan *et al*., 1989; Jagannathan, 1990, 1993; Khan and Jagannathan, 1995; Jagannathan and Khan, 1996, 1997, 1998; Conte *et al*., 1996; Khan, 1997), both in the scalar and the spinor cases, respectively. A very detailed description of these transforms and techniques is available in (Jagannathan and Khan, 1996).

Now, we can compare the aforementioned Hamiltonians with the conventional Hamiltonian given by the square-root approach (Dragt, 1998). The square-root approach does not give all the terms, such as the one involving the commutator of $\mathbf{p}_\perp^2$ with $\frac{\partial}{\partial z}\left(n^2(\mathbf{r})\right)$. Our procedure of linearization and expansion in powers of $|\widehat{\mathbf{p}}_\perp|/n_0$ gives all the terms, which one gets by the square-root expansion of (3) and some additional terms, which are the wavelength-dependent terms. Such wavelength-dependent terms can in no way be obtained by any of the conventional prescriptions, starting with the Helmholtz equation (6).

4. APPLICATIONS

In the previous sections, we presented an alternative to the square-root expansion and obtained an expansion for the beam-optical Hamiltonian, which works to all orders. Formal expressions were obtained for the paraxial Hamiltonian and the leading order aberrating Hamiltonian, without assuming any form for the refractive index. Even at the paraxial level the wavelength-dependent effects manifest by the presence of a commutator term, which does not vanish for a varying refractive index.

Now, we apply the formalism to specific examples. First one is the medium with a constant refractive index. This is perhaps the only problem, which can be solved exactly in a closed form expression. This example is just to illustrate how the aberration expansion in our formalism can be summed to give the familiar exact result.

The next example is that of the axially symmetric graded-index medium. This example enables us to demonstrate the power of the formalism, reproducing the familiar results from the traditional approaches and further giving rise to new results, dependent on the wavelength.

4.1. Medium with Constant Refractive Index

Constant refractive index is the simplest possible system. In our formalism, this is perhaps the only case where it is possible to do an exact diagonalization. This is very similar to the exact diagonalization of the free Dirac Hamiltonian. From the experience of the Dirac theory we know that there are hardly any situations, where one can do the exact diagonalization. One necessarily has to resort to some approximate diagonalization procedure. The Foldy–Wouthuysen transformation scheme provides the most convenient and accurate diagonalization to any desired degree of accuracy. So, we have adopted the Foldy–Wouthuysen scheme in our formalism.

For a medium with constant refractive index, $n(\mathbf{r}) = n_c$, we have,

$$\widehat{\mathrm{H}}_c = -n_0\sigma_z + \mathcal{D}\sigma_z + \mathcal{D}(i\sigma_y)$$
$$\mathcal{D} = \frac{1}{2n_0}\left\{\widehat{\mathbf{p}}_\perp^2 + \left(n_0^2 - n_c^2\right)\right\}. \tag{24}$$

The Hamiltonian in (24) can be exactly diagonalized by the following transformation,

$$\begin{aligned} T^{\pm} &= \exp\left[i(\pm i\sigma_z)\widehat{\mathcal{O}}\theta\right] \\ &= \exp\left[\mp\sigma_x\mathcal{D}\theta\right] \\ &= \cosh\left(\mathcal{D}\theta\right) \mp \sigma_x \sinh\left(\mathcal{D}\theta\right). \end{aligned} \tag{25}$$

We choose,

$$\tanh\left(2\mathcal{D}\theta\right) = \frac{\mathcal{D}}{n_0 - \mathcal{D}} = \frac{n_0^2 - \left(n_c^2 - \widehat{\mathbf{p}}_\perp^2\right)}{n_0^2 + \left(n_c^2 + \widehat{\mathbf{p}}_\perp^2\right)} < 1, \tag{26}$$

then we obtain,

$$\begin{aligned} \widehat{\mathrm{H}}_c^{\text{diagonal}} &= T^+\widehat{\mathrm{H}}_c T^- \\ &= T^+\{-n_0\sigma_z + \mathcal{D}\sigma_z + \mathcal{D}(i\sigma_y)\}T^- \\ &= -\sigma_z\left\{n_0^2 - 2n_0\mathcal{D}\right\}^{\frac{1}{2}} \\ &= -\sigma_z\left\{n_c^2 - \widehat{\mathbf{p}}_\perp^2\right\}^{\frac{1}{2}}. \end{aligned} \tag{27}$$

We next, compare the exact result thus obtained with the approximate one, obtained through the systematic series procedure we have developed. We define $P = \frac{1}{n_0^2}\{\widehat{\mathbf{p}}_\perp^2 + (n_0^2 - n_c^2)\}$. Then,

$$\begin{aligned} \widehat{\mathcal{H}}_c^{(4)} &= -n_0\left\{1 - \frac{1}{2}P - \frac{1}{8}P^2 - \frac{1}{16}P^3 - \frac{5}{128}P^4\right\}\sigma_z \\ &\approx -n_0\{1 - P^2\}^{\frac{1}{2}} \\ &= -\left\{n_c^2 - \widehat{\mathbf{p}}_\perp^2\right\}^{\frac{1}{2}} \\ &= \widehat{\mathrm{H}}_c^{\text{diagonal}}. \end{aligned} \tag{28}$$

Knowing the Hamiltonian, we can compute the transfer maps. The transfer operator between any pair of points $\{(z'', z')|z'' > z'\}$ on the z-axis, is formally given by

$$|\psi(z'', z')| = \widehat{\mathcal{T}}(z'', z')|\psi(z'', z')\rangle, \tag{29}$$

with

$$
\begin{aligned}
&i\lambda\!\!\!^{-}\frac{\partial}{\partial z}\widehat{\mathcal{T}}(z'', z') = \widehat{\mathcal{H}}\widehat{\mathcal{T}}(z'', z'), \quad \widehat{\mathcal{T}}(z'', z') = \widehat{\mathcal{I}}, \\
&\qquad \widehat{\mathcal{T}}(z'', z') = \wp\left\{\exp\left[-\frac{i}{\lambda\!\!\!^{-}}\int_{z'}^{z''} dz\,\widehat{\mathcal{H}}(z)\right]\right\} \\
&\qquad\qquad = \widehat{\mathcal{I}} - \frac{i}{\lambda\!\!\!^{-}}\int_{z'}^{z''} dz\widehat{\mathcal{H}}(z) + \left(-\frac{i}{\lambda\!\!\!^{-}}\right)^2 \int_{z'}^{z''} dz \int_{z'}^{z} dz'\widehat{\mathcal{H}}(z)\widehat{\mathcal{H}}(z') + \cdots,
\end{aligned} \tag{30}
$$

where $\widehat{\mathcal{I}}$ is the identity operator and $\wp$ denotes the path-ordered exponential. There is no closed form expression for $\widehat{\mathcal{T}}(z'', z')$ for an arbitrary choice of the refractive index $n(\mathbf{r})$. In such a situation, the most convenient form of the expression for the z-evolution operator $\widehat{\mathcal{T}}(z'', z')$, or the z-propagator, is

$$
\widehat{\mathcal{T}}(z'', z') = \exp\left[-\frac{i}{\lambda\!\!\!^{-}}\widehat{T}(z'', z')\right], \tag{31}
$$

with

$$
\hat{T}(z'', z') = \int_{z'}^{z''} dz\widehat{\mathcal{H}}(z) + \frac{1}{2}\left(-\frac{i}{\lambda\!\!\!^{-}}\right)\int_{z'}^{z''} dz \int_{z'}^{z} dz'[\widehat{\mathcal{H}}(z), \widehat{\mathcal{H}}(z')] + \cdots, \tag{32}
$$

as given by the Magnus (1954) formula, which is described in detail in Appendix C. We shall be needing these expressions in the next example where the refractive index is not a constant.

Using the procedure outlined earlier, we compute the transfer operator,

$$
\begin{aligned}
\widehat{U}_c\,(z_{\text{out}}, z_{\text{in}}) &= \exp\left[-\frac{i}{\lambda\!\!\!^{-}}\Delta z \mathcal{H}_c\right] \\
&= \exp\left[+\frac{i}{\lambda\!\!\!^{-}} n_c \Delta z\left\{1 - \frac{1}{2}\frac{\widehat{p}_\perp^2}{n_c^2} - \frac{1}{8}\left(\frac{\widehat{p}_\perp^2}{n_c^2}\right)^2 - \cdots\right\}\right], \\
\Delta z &= z_{\text{out}} - z_{\text{in}},
\end{aligned} \tag{33}
$$

where 'in' and 'out' are the 'input' and 'output' plains along the optic-axis at the points z_{in} and z_{out}, respectively. Using (33), we compute the transfer maps

$$
\begin{pmatrix} \langle \mathbf{r}_\perp \rangle \\ \langle \mathbf{p}_\perp \rangle \end{pmatrix}_{\text{out}} = \begin{pmatrix} 1 & \frac{1}{\sqrt{n_c^2 - \mathbf{p}_\perp^2}}\Delta z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \langle \mathbf{r}_\perp \rangle \\ \langle \mathbf{p}_\perp \rangle \end{pmatrix}_{\text{in}}. \tag{34}
$$

The beam-optical Hamiltonian is intrinsically aberrating. Even for the simplest situation of a constant refractive index, we have aberrations to all orders!

4.2. Axially Symmetric Graded-Index Medium

We just saw the treatment of the medium with a constant refractive index. This is perhaps the only problem which can be solved exactly in a closed form expression. This example was just to illustrate how the aberration expansion in our formalism can be obtained. We now consider the next example. The refractive index of an axially symmetric graded-index material can be most generally described by the following polynomial (see, pp. 117 in Dragt *et al.*, 1986)

$$n(\mathbf{r}) = n_0 + \alpha_2(z)\mathbf{r}_\perp^2 + \alpha_4(z)\mathbf{r}_\perp^4 + \alpha_6(z)\mathbf{r}_\perp^6 + \alpha_8(z)\mathbf{r}_\perp^8 + \cdots , \tag{35}$$

where we have assumed the axis of symmetry to coincide with the optic-axis, namely the z-axis without any loss of generality. To write the beam-optical Hamiltonians we introduce the following notation

$$\begin{aligned}
\widehat{T} &= (\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp) \\
w_1(z) &= \frac{d}{dz}\{2n_0\alpha_2(z)\} \\
w_2(z) &= \frac{d}{dz}\left\{\alpha_2^2(z) + 2n_0\alpha_4(z)\right\} \\
w_3(z) &= \frac{d}{dz}\{2n_0\alpha_6(z) + 2\alpha_2(z)\alpha_4(z)\} \\
w_4(z) &= \frac{d}{dz}\left\{\alpha_4^2(z) + 2\alpha_2(z)\alpha_6(z) + 2n_0\alpha_8(z)\right\} .
\end{aligned} \tag{36}$$

We also use, $[A, B]_+ = (AB + BA)$. The beam-optical Hamiltonian is

$$\widehat{\mathcal{H}} = \widehat{H}_{0,p} + \widehat{H}_{0,(4)} + \widehat{H}_{0,(6)} + \widehat{H}_{0,(8)} + \widehat{H}_{0,(2)}^{(\lambda\!\!\!^-)} + \widehat{H}_{0,(4)}^{(\lambda\!\!\!^-)} + \widehat{H}_{0,(6)}^{(\lambda\!\!\!^-)} + \widehat{H}_{0,(8)}^{(\lambda\!\!\!^-)}$$

$$\begin{aligned}
\widehat{H}_{0,p} &= -n_0 + \frac{1}{2n_0}\widehat{\mathbf{p}}_\perp^2 - \alpha_2(z)\mathbf{r}_\perp^2 \\
\widehat{H}_{0,(4)} &= \frac{1}{8n_0^3}\widehat{\mathbf{p}}_\perp^4 - \frac{\alpha_2(z)}{4n_0^2}(\widehat{\mathbf{p}}_\perp^2\mathbf{r}_\perp^2 + \mathbf{r}_\perp^2\widehat{\mathbf{p}}_\perp^2) - \alpha_4(z)\mathbf{r}_\perp^4 \\
\widehat{H}_{0,(6)} &= \frac{1}{16n_0^5}\widehat{\mathbf{p}}_\perp^6 - \frac{\alpha_2(z)}{8n_0^4}\{(\widehat{\mathbf{p}}_\perp^4\mathbf{r}_\perp^2 + \mathbf{r}_\perp^2\widehat{\mathbf{p}}_\perp^4) + \widehat{\mathbf{p}}_\perp^2\mathbf{r}_\perp^2\widehat{\mathbf{p}}_\perp^2\} \\
&\quad + \frac{1}{8n_0^3}\left\{\left(\alpha_2^2(z) - 2n_0\alpha_4(z)\right)(\widehat{\mathbf{p}}_\perp^2\mathbf{r}_\perp^4 + \mathbf{r}_\perp^4\widehat{\mathbf{p}}_\perp^2) + 2\alpha_2^2(z)\mathbf{r}_\perp^2\widehat{\mathbf{p}}_\perp^2\mathbf{r}_\perp^2\right\} - \alpha_6(z)\mathbf{r}_\perp^6 \\
\widehat{H}_{0,(8)} &= \frac{5}{128n_0^7}\widehat{\mathbf{p}}_\perp^8 - \frac{5\alpha_2(z)}{64n_0^6}[\widehat{\mathbf{p}}_\perp^4, [\widehat{\mathbf{p}}_\perp^2\mathbf{r}_\perp^2]_+]_+ \\
&\quad + \frac{1}{32n_0^5}\left\{ \left(3\alpha_2^2(z) - 4n_0\alpha_4(z)\right)[\widehat{\mathbf{p}}_\perp^4, \mathbf{r}_\perp^4]_+ + 5\alpha_2^2(z)[\widehat{\mathbf{p}}_\perp^2, \mathbf{r}_\perp^2]_+^2\right.
\end{aligned}$$

$$
\begin{aligned}
& - \left(2\alpha_2^2(z) + 4n_0\alpha_4(z)\right) \widehat{\mathbf{p}}_\perp^2 \mathbf{r}_\perp^4 \widehat{\mathbf{p}}_\perp^2 \big\} \\
& + \frac{1}{16n_0^4} \big\{ 4\left(\alpha_2^3(z) + n_0\alpha_2(z)\alpha_4(z) + n_0^2\alpha_6(z)\right) [\widehat{\mathbf{p}}_\perp^2, \mathbf{r}_\perp^6]_+ \\
& - 5\alpha_2^3(z)[\mathbf{r}_\perp^4, [\widehat{\mathbf{p}}_\perp^2, \mathbf{r}_\perp^2]_+]_+ \\
& + \left(2\alpha_2^3(z) + 4n_0\alpha_2(z)\alpha_4(z)\right) [\mathbf{r}_\perp^2, \mathbf{r}_\perp^2 \widehat{\mathbf{p}}_\perp^2 \mathbf{r}_\perp^2]_+ \big\} \\
& - \alpha_8(z)\mathbf{r}_\perp^8 \\
\widehat{H}_{0,(2)}^{(\lambda\!\!\!^-)} &= -\frac{\lambda\!\!\!^-{}^2}{16n_0^4} \left\{ \frac{d}{dz}\left(n_0\alpha_2(z)\right) \right\} \widehat{T} \\
\widehat{H}_{0,(4)}^{(\lambda\!\!\!^-)} &= -\frac{\lambda\!\!\!^-{}^2}{32n_0^4} w_2(z)(\mathbf{r}_\perp^2 \widehat{T} + \widehat{T}\mathbf{r}_\perp^2) + \frac{\lambda\!\!\!^-{}^2}{32n_0^5} w_1^2(z)\mathbf{r}_\perp^4 \\
\widehat{H}_{0,(6)}^{(\lambda\!\!\!^-)} &= -\frac{3\lambda\!\!\!^-{}^2}{32n_0^4} w_3(z)(\mathbf{r}_\perp^4 \widehat{T} + \widehat{T}\mathbf{r}_\perp^4) + \frac{\lambda\!\!\!^-{}^2}{16n_0^5} w_1(z)w_2(z)\mathbf{r}_\perp^6 \\
\widehat{H}_{0,(8)}^{(\lambda\!\!\!^-)} &= -\frac{\lambda\!\!\!^-{}^2}{8n_0^4} w_4(z)(\mathbf{r}_\perp^6 \widehat{T} + \widehat{T}\mathbf{r}_\perp^6) \\
& + \frac{\lambda\!\!\!^-{}^2}{32n_0^5} \left\{ w_2^2(z) + 2w_1(z)w_3(z) \right\} \mathbf{r}_\perp^8 . \qquad (37)
\end{aligned}
$$

The reason for partitioning $\widehat{\mathcal{H}}$ in the aforementioned manner will be clear as we proceed.

4.2.1. The Paraxial Hamiltonian

For this system, the commutator term modifies the paraxial Hamiltonian as

$$
\begin{aligned}
\widehat{\mathcal{H}}^{(2)} &= \widehat{H}_{0,p} + \widehat{H}_{0,(2)}^{(\lambda\!\!\!^-)} \\
&= -n_0 + \frac{1}{2n_0} \widehat{\mathbf{p}}_\perp^2 - \alpha_2(z)\mathbf{r}_\perp^2 - \frac{\lambda\!\!\!^-{}^2}{16n_0^3} \left\{ \frac{d}{dz}\alpha_2(z) \right\} \left(\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp\right) .
\end{aligned} \qquad (38)
$$

The extra term in the aforementioned equation is bound to have a bearing on the beam optics of the system.

The paraxial transfer maps are formally given by

$$
\begin{pmatrix} \langle \mathbf{r}_\perp \rangle \\ \langle \mathbf{p}_\perp \rangle \end{pmatrix}_{\text{out}} = \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} \langle \mathbf{r}_\perp \rangle \\ \langle \mathbf{p}_\perp \rangle \end{pmatrix}_{\text{in}} , \qquad (39)
$$

where P, Q, R and S are the solutions of the paraxial Hamiltonian in (37). The symplecticity condition tells us that $PS - QR = 1$. In this particular case from the structure of the paraxial equations, we can further conclude that: $R = P'$ and $S = Q'$ where $(\,)'$ denotes the z-derivative.

4.2.2. Aberrations

The Hamiltonian $\widehat{H}_{0,(4)}$ is the one we have in the traditional prescriptions and is responsible for the six aberrations. $\widehat{H}_{0,(4)}^{(\lambda\!\!\!^{-})}$ modifies the six aberrations mentioned earlier by wavelength-dependent contributions. These six aberrations are:

Symbol	Polynomial	Name
C	$\widehat{\mathbf{p}}_\perp^4$	Spherical aberration
K	$[\widehat{\mathbf{p}}_\perp^2, (\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp)]_+$	Coma
A	$(\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp)^2$	Astigmatism
F	$(\widehat{\mathbf{p}}_\perp^2 \mathbf{r}_\perp^2 + \mathbf{r}_\perp^2 \widehat{\mathbf{p}}_\perp^2)$	Curvature of field
D	$[\mathbf{r}_\perp^2, (\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp)]_+$	Distortion
E	$\mathbf{r}_\perp^4$	Nameless? or POCUS

The name *POCUS* is used in (Dragt *et al.*, 1986) on page 137.

The transfer operator is most accurately expressed in terms of the paraxial solutions, P, Q, R and S, *via* the *interaction picture* (Dragt and Forest, 1986).

$$
\begin{aligned}
\widehat{\mathcal{T}}(z, z_0) &= \exp\left[-\frac{i}{\lambda\!\!\!^{-}}\widehat{T}(z, z_0)\right], \\
&= \exp\left[-\frac{i}{\lambda\!\!\!^{-}}\left\{C(z'', z')\widehat{\mathbf{p}}_\perp^4\right.\right. \\
&\quad + K(z'', z')[\widehat{\mathbf{p}}_\perp^2, (\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp)]_+ \\
&\quad + A(z'', z')\,(\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp)^2 \\
&\quad + F(z'', z')(\widehat{\mathbf{p}}_\perp^2 \mathbf{r}_\perp^2 + \mathbf{r}_\perp^2 \widehat{\mathbf{p}}_\perp^2) \\
&\quad + D(z'', z')[\mathbf{r}_\perp^2, (\widehat{\mathbf{p}}_\perp \cdot \mathbf{r}_\perp + \mathbf{r}_\perp \cdot \widehat{\mathbf{p}}_\perp)]_+ \\
&\quad \left.\left. + E(z'', z')\mathbf{r}_\perp^4\right\}\right].
\end{aligned}
\tag{40}
$$

The six aberration coefficients are given by,

$$
C(z'', z') = \int_{z'}^{z''} dz \left\{ \frac{1}{8n_0^3} S^4 - \frac{\alpha_2(z)}{2n_0^2} Q^2 S^2 - \alpha_4(z) Q^4 - \frac{\lambda\!\!\!^{-2}}{8n_0^4} w_2(z) Q^3 S \right.
$$

$$
\begin{aligned}
&\quad + \frac{\lambdabar^2}{32 n_0^5} w_1^2(z) Q^4 \bigg\}, \\
K(z'', z') &= \int_{z'}^{z''} dz \left\{ \frac{1}{8 n_0^3} R S^3 - \frac{\alpha_2(z)}{4 n_0^2} Q S (PS + QR) - \alpha_4(z) P Q^3 \right. \\
&\quad - \frac{\lambdabar^2}{32 n_0^4} w_2(z) (Q^2 (PS + QR) + 2 P Q^2 S) \\
&\quad \left. + \frac{\lambdabar^2}{32 n_0^5} w_1^2(z) P Q^3 \right\}, \\
A(z'', z') &= \int_{z'}^{z''} dz \left\{ \frac{1}{8 n_0^3} R^2 S^2 - \frac{\alpha_2(z)}{2 n_0^2} P Q R S - \alpha_4(z) P^2 Q^2 \right. \\
&\quad - \frac{\lambdabar^2}{16 n_0^4} w_2(z) (P Q (PS + QR)) \\
&\quad \left. + \frac{\lambdabar^2}{32 n_0^5} w_1^2(z) P^2 Q^2 \right\}, \\
F(z'', z') &= \int_{z'}^{z''} dz \left\{ \frac{1}{8 n_0^3} R^2 S^2 - \frac{\alpha_2(z)}{4 n_0^2} (P^2 S^2 + Q^2 R^2) - \alpha_4(z) P^2 Q^2 \right. \\
&\quad - \frac{\lambdabar^2}{16 n_0^4} w_2(z) (P Q (PS + QR)) \\
&\quad \left. + \frac{\lambdabar^2}{32 n_0^5} w_1^2(z) P^2 Q^2 \right\}, \\
D(z'', z') &= \int_{z'}^{z''} dz \left\{ \frac{1}{8 n_0^3} R^3 S - \frac{\alpha_2(z)}{4 n_0^2} P R (PS + QR) - \alpha_4(z) P^3 Q \right. \\
&\quad - \frac{\lambdabar^2}{32 n_0^4} w_2(z) (P^2 (PS + QR) + 2 P^2 Q R) \\
&\quad \left. + \frac{\lambdabar^2}{32 n_0^5} w_1^2(z) P^3 Q \right\}, \\
E(z'', z') &= \int_{z'}^{z''} dz \left\{ \frac{1}{8 n_0^3} R^4 - \frac{\alpha_2(z)}{2 n_0^2} P^2 R^2 - \alpha_4(z) P^4 \right. \\
&\quad - \frac{\lambdabar^2}{8 n_0^4} w_2(z) (P^3 R)
\end{aligned}
$$

$$+\frac{\lambda\!\!\!^{-}{}^2}{32n_0^5}w_1^2(z)P^4\bigg\}. \tag{41}$$

Thus, we see that the transfer operator and the aberration coefficients are modified by $\lambda\!\!\!^{-}$-dependent contributions. The traditional beam optics is obtained in the limit $\lambda\!\!\!^{-} \longrightarrow 0$. The sixth- and eighth-order Hamiltonians are modified by the presence of wavelength-dependent terms. These will in turn modify the fifth- and seventh-order aberrations, respectively.

5. CONCLUDING REMARKS

We exploited the similarities between the Helmholtz equation and the Klein–Gordon equation to obtain an alternate prescription for the aberration expansion. In this prescription, we followed a procedure due to Feshbach–Villars for linearizing the Klein–Gordon equation. After casting the Helmholtz equation to this linear form, it was further possible to use the Foldy–Wouthuysen transformation technique of the Dirac electron theory. This enabled us to obtain the beam-optical Hamiltonian to any desired degree of accuracy. We further get the wavelength-dependent contributions at each order, starting with the lowest-order paraxial Hamiltonian. Formal expressions were obtained for the paraxial and leading order aberrating Hamiltonians, without making any assumptions on the form of the refractive index.

It is interesting that the extra commutator term $-\frac{i\lambda\!\!\!^{-}}{32n_0^4}[\widehat{\mathbf{p}}_\perp^2, \frac{\partial}{\partial z}(n^2(\mathbf{r}))]$ in (22) contributes a correction to the optical Hamiltonian, even at the 'paraxial-level,' when the refractive index of the medium suffers both longitudinal and transverse inhomogeneities. Such a z-derivative term is not natural to the traditional power series expansion. This commutator term originates from $-\frac{1}{8n_0^2}[\widehat{\mathcal{O}}, ([\widehat{\mathcal{O}}, \widehat{\mathcal{E}}] + i\lambda\!\!\!^{-}\frac{\partial}{\partial z}\widehat{\mathcal{O}})]$ in the expression for $\widehat{\mathcal{H}}^{(4)}$ in (22). In the Foldy–Wouthuysen formalism of the Dirac theory the corresponding commutator term is responsible for the correct explanation of the spin-orbit energy (including the Thomas precession effect) and the Darwin term (attributed to the zitterbewegung) (see Section 4.3 of Bjorken and Drell, 1964). Similarly, in the nonrelativistic reduction and interpretation of the Klein–Gordon equation using the Foldy–Wouthuysen transformation theory such a commutator term corresponds to the Darwin term correcting the classical electrostatic interaction of a point charge in analogy to the zitterbewegung of the Dirac electron (see Section 9.7 of Bjorken and Drell, 1964). In the quantum theory of beam optics of charged Klein–Gordon and Dirac particles (Khan and Jagannathan, 1995; Jagannathan and Khan, 1996, 1997, 1998; Conte *et al.*, 1996; Khan, 1997) the corresponding terms add to the Hamiltonian the lowest order quantum corrections to the classical aberration terms. In view of this analogy, it should be of interest to study the effect of this correction term to the optical Hamiltonian.

As an example, we considered the *medium with a constant refractive index*. This is perhaps the only problem, which can be solved exactly, in a closed form expression. This example was primarily for illustrating certain aspects of the machinery we have used. The second, and the more interesting example is that of the *axially symmetric graded-index medium*. For this system, we derived the beam-optical Hamiltonians to eighth order. At each order, we find the wavelength-dependent contributions. The fourth-order Hamiltonian was used to obtain the six, third-order aberrations coefficients which get modified by the wavelength-dependent contributions. Explicit relations for these coefficients were presented. In the limit $\lambdabar \longrightarrow 0$, the alternate prescription here, reproduces the very well known *Lie Algebraic Formalism of Light Optics* (Dragt, 1998). It would be worthwhile to look for the extra wavelength-dependent contributions experimentally.

The close analogy between geometrical optics and charged-particle has been known for too long a time. Until recently, it was possible to see this analogy only between the geometrical optics and classical prescriptions of charged-particle optics. A quantum theory of charged-particle optics was presented in recent years (Jagannathan *et al.*, 1989; Jagannathan, 1990, 1993; Khan and Jagannathan, 1995; Jagannathan and Khan, 1996, 1997, 1998; Conte *et al.*, 1996; Khan, 1997). With the current development of the non-traditional prescriptions of Helmholtz optics (Khan *et al.*, 2002) and the matrix formulation of Maxwell optics, using the rich algebraic machinery of quantum mechanics, it is now possible to see a parallel of the analogy at each level. The non-traditional prescription of the Helmholtz optics is in close analogy with the quantum theory of charged-particles based on the Klein–Gordon equation (Khan, 2002).

APPENDIX A: THE FESHBACH—VILLARS FORM OF THE KLEIN—GORDON EQUATION

The method we have followed to cast the time-independent Klein–Gordon equation into a beam-optical form linear in $\frac{\partial}{\partial z}$, suitable for a systematic study, through successive approximations, using the Foldy–Wouthuysen-like transformation technique borrowed from the Dirac theory, is similar to the way the time-dependent Klein–Gordon equation is transformed (Feshbach and Villars, 1958) to the Schrödinger form, containing only first-order time derivative, in order to study its nonrelativistic limit using the Foldy–Wouthuysen technique (see, *e.g.*, Bjorken and Drell, 1964).

Defining

$$\Phi = \frac{\partial}{\partial t}\Psi, \tag{A.1}$$

the free particle Klein–Gordon equation is written as

$$\frac{\partial}{\partial t}\Phi = \left(c^2\nabla^2 - \frac{{m_0}^2c^4}{\hbar^2}\right)\Psi. \tag{A.2}$$

Introducing the linear combinations

$$\Psi_+ = \frac{1}{2}\left(\Psi + \frac{i\hbar}{m_0c^2}\Phi\right), \qquad \Psi_- = \frac{1}{2}\left(\Psi - \frac{i\hbar}{m_0c^2}\Phi\right) \tag{A.3}$$

the Klein–Gordon equation is seen to be equivalent to a pair of coupled differential equations:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Psi_+ &= -\frac{\hbar^2\nabla^2}{2m_0}\left(\Psi_+ + \Psi_-\right) + m_0c^2\Psi_+ \\ i\hbar\frac{\partial}{\partial t}\Psi_- &= \frac{\hbar^2\nabla^2}{2m_0}\left(\Psi_+ + \Psi_-\right) - m_0c^2\Psi_-. \end{aligned} \tag{A.4}$$

Equation (A.4) can be written in a two-component language as

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} = \widehat{H}_0^{\mathrm{FV}}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix}, \tag{A.5}$$

with the Feshbach–Villars Hamiltonian for the free particle, $\widehat{H}_0^{\mathrm{FV}}$, given by

$$\widehat{H}_0^{\mathrm{FV}} = \begin{pmatrix} m_0c^2 + \frac{\widehat{p}^2}{2m_0} & \frac{\widehat{p}^2}{2m_0} \\ -\frac{\widehat{p}^2}{2m_0} & -m_0c^2 - \frac{\widehat{p}^2}{2m_0}\end{pmatrix} = m_0c^2\sigma_z + \frac{\widehat{p}^2}{2m_0}\sigma_z + i\frac{\widehat{p}^2}{2m_0}\sigma_y. \tag{A.6}$$

For a free nonrelativistic particle with kinetic energy $\ll m_0c^2$, it is seen that Ψ_+ is large compared to Ψ_-.

In presence of an electromagnetic field, the interaction is introduced through the minimal coupling

$$\widehat{\mathbf{p}} \longrightarrow \widehat{\boldsymbol{\pi}} = \widehat{\mathbf{p}} - q\mathbf{A}, \qquad i\hbar\frac{\partial}{\partial t} \longrightarrow i\hbar\frac{\partial}{\partial t} - q\phi. \tag{A.7}$$

The corresponding Feshbach–Villars form of the Klein–Gordon equation becomes

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} &= \widehat{H}^{\mathrm{FV}}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} \\ \begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} &= \frac{1}{2}\begin{pmatrix}\Psi + \frac{1}{m_0c^2}\left(i\hbar\frac{\partial}{\partial t} - q\phi\right)\Psi \\ \Psi - \frac{1}{m_0c^2}\left(i\hbar\frac{\partial}{\partial t} - q\phi\right)\Psi\end{pmatrix} \\ \widehat{H}^{\mathrm{FV}} &= m_0c^2\sigma_z + \widehat{\mathcal{E}} + \widehat{\mathcal{O}} \\ \widehat{\mathcal{E}} &= q\phi + \frac{\widehat{\pi}^2}{2m_0}\sigma_z, \qquad \widehat{\mathcal{O}} = i\frac{\widehat{\pi}^2}{2m_0}\sigma_y. \end{aligned} \tag{A.8}$$

As in the free-particle case, in the nonrelativistic situation Ψ_+ is large compared to Ψ_-. The even term $\widehat{\mathcal{E}}$ does not couple Ψ_+ and Ψ_-, whereas $\widehat{\mathcal{O}}$ is odd which couples Ψ_+ and Ψ_-. Starting from (A.8), the nonrelativistic limit of the Klein–Gordon equation, with various correction terms, can be understood using the Foldy–Wouthuysen technique (see *e.g.*, Bjorken and Drell, 1964).

It is clear from the aforementioned discussion that we have just adopted the technique mentioned earlier for studying the z-evolution of the Klein–Gordon wave function of a charged-particle beam in an optical system comprising a static electromagnetic field. The additional feature of our formalism is the extra approximation of dropping σ_z in an intermediate stage to take into account the fact that we are interested only in the forward-propagating beam along the z-direction.

APPENDIX B: FOLDY–WOUTHUYSEN TRANSFORMATION

In the traditional scheme, the purpose of expanding the *light optics* Hamiltonian $\widehat{H} = -(n^2(\mathbf{r}) - \widehat{\mathbf{p}}_\perp^2)^{1/2}$ in a series using $(\frac{1}{n_0^2}\widehat{\mathbf{p}}_\perp^2)$ as the expansion parameter is to understand the propagation of the quasi-paraxial beam in terms of a series of approximations (paraxial + nonparaxial). Similar is the situation in the case of the *charged-particle optics*. Let us recall that in relativistic quantum mechanics too one has a similar problem of understanding the relativistic wave equations as the nonrelativistic approximation plus the relativistic correction terms in the quasi-relativistic regime. For the Dirac equation (which is first order in time), this is done most conveniently using the Foldy–Wouthuysen transformation leading to an iterative diagonalization technique.

The main framework of the formalism of optics, used here (and in the charged-particle optics) is based on the transformation technique of the Foldy–Wouthuysen theory which casts the Dirac equation in a form displaying the different interaction terms between the Dirac particle and an applied electromagnetic field in a nonrelativistic and easily interpretable form (see Foldy and Wouthuysen, 1950; Pryce, 1948; Tani, 1951; Acharya and Sudarshan, 1960, for a general discussion of the role of the Foldy–Wouthuysen-type transformations in particle interpretation of relativistic wave equations). In the Foldy–Wouthuysen theory, the Dirac equation is decoupled through a canonical transformation into two two-component equations: one reduces to the Pauli equation in the nonrelativistic limit and the other describes the negative-energy states.

Let us describe here briefly the standard Foldy–Wouthuysen theory so that the way it has been adopted for the purposes of the earlier studies in optics will be clear. Let us consider a charged-particle of rest-mass m_0, charge q in the presence of an electromagnetic field characterized by $\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial}{\partial t}\boldsymbol{A}$ and $\mathbf{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$. Then the Dirac equation is

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, t) = \widehat{H}_{\mathrm{D}}\Psi(\mathbf{r}, t) \qquad \text{(B.1)}$$

$$\begin{aligned}
\widehat{H}_{\mathrm{D}} &= m_0c^2\beta + q\phi + c\boldsymbol{\alpha}\cdot\widehat{\boldsymbol{\pi}} \\
&= m_0c^2\beta + \widehat{\mathcal{E}} + \widehat{\mathcal{O}} \\
\widehat{\mathcal{E}} &= q\phi \\
\widehat{\mathcal{O}} &= c\boldsymbol{\alpha}\cdot\widehat{\boldsymbol{\pi}},
\end{aligned} \tag{B.2}$$

where

$$\begin{aligned}
&\boldsymbol{\alpha} = \begin{bmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{bmatrix}, \qquad \beta = \begin{bmatrix} \mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\mathbb{1} \end{bmatrix}, \qquad \mathbb{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\
&\boldsymbol{\sigma} = \left[\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right].
\end{aligned} \tag{B.3}$$

with $\widehat{\boldsymbol{\pi}} = \widehat{\boldsymbol{p}} - q\boldsymbol{A}$, $\widehat{\boldsymbol{p}} = -i\hbar\boldsymbol{\nabla}$, and $\widehat{\pi}^2 = \left(\widehat{\pi}_x^2 + \widehat{\pi}_y^2 + \widehat{\pi}_z^2\right)$.

In the nonrelativistic situation, the upper pair of components of the Dirac Spinor Ψ are large compared to the lower pair of components. The operator $\widehat{\mathcal{E}}$ which does not couple the large and small components of Ψ is called 'even' and $\widehat{\mathcal{O}}$ is called an 'odd' operator, which couples the large to the small components. Note that

$$\beta\widehat{\mathcal{O}} = -\widehat{\mathcal{O}}\beta, \qquad \beta\widehat{\mathcal{E}} = \widehat{\mathcal{E}}\beta. \tag{B.4}$$

Now, the search is for a unitary transformation, $\Psi' = \Psi \longrightarrow \widehat{U}\Psi$, such that the equation for Ψ' does not contain any odd operator.

In the free particle case (with $\phi = 0$ and $\widehat{\boldsymbol{\pi}} = \widehat{\mathbf{p}}$), such a Foldy–Wouthuysen transformation is given by

$$\begin{aligned}
&\Psi \longrightarrow \Psi' = \widehat{U}_{\mathrm{F}}\Psi \\
&\widehat{U}_{\mathrm{F}} = e^{i\widehat{S}} = e^{\beta\boldsymbol{\alpha}\cdot\widehat{\mathbf{p}}\theta}, \qquad \tan 2|\widehat{\mathbf{p}}|\theta = \frac{|\widehat{\mathbf{p}}|}{m_0c}.
\end{aligned} \tag{B.5}$$

This transformation eliminates the odd part completely from the free particle Dirac Hamiltonian, reducing it to the diagonal form:

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}\Psi' &= e^{i\widehat{S}}(m_0c^2\beta + c\boldsymbol{\alpha}\cdot\widehat{\mathbf{p}})e^{-i\widehat{S}}\Psi' \\
&= \left(\cos|\widehat{\mathbf{p}}|\theta + \frac{\beta\boldsymbol{\alpha}\cdot\widehat{\mathbf{p}}}{|\widehat{\mathbf{p}}|}\sin|\widehat{\mathbf{p}}|\theta\right)(m_0c^2\beta + c\boldsymbol{\alpha}\cdot\widehat{\mathbf{p}}) \\
&\quad \times \left(\cos|\widehat{\mathbf{p}}|\theta - \frac{\beta\boldsymbol{\alpha}\cdot\widehat{\mathbf{p}}}{|\widehat{\mathbf{p}}|}\sin|\widehat{\mathbf{p}}|\theta\right)\Psi' \\
&= (m_0c^2\cos 2|\widehat{\mathbf{p}}|\theta + c|\widehat{\mathbf{p}}|\sin 2|\widehat{\mathbf{p}}|\theta)\beta\Psi' \\
&= \left(\sqrt{m_0^2c^4 + c^2\widehat{p}^2}\right)\beta\,\Psi'.
\end{aligned} \tag{B.6}$$

In the general case, when the electron is in a time-dependent electromagnetic field it is not possible to construct an $\exp(i\widehat{S})$, which removes the odd operators from the transformed Hamiltonian completely. Therefore, one has to be content with a nonrelativistic expansion of the transformed Hamiltonian in a power series in $1/m_0c^2$ keeping through any desired order. Note that in the nonrelativistic case, when $|\mathbf{p}| \ll m_0c$, the transformation operator $\widehat{U}_{\mathrm{F}} = \exp(i\widehat{S})$ with $\widehat{S} \approx -i\beta\widehat{\mathcal{O}}/2m_0c^2$, where $\widehat{\mathcal{O}} = c\boldsymbol{\alpha}\cdot\widehat{\mathbf{p}}$ is the odd part of the free Hamiltonian. So, in the general case, we can start with the transformation

$$\Psi^{(1)} = e^{i\widehat{S}_1}\Psi, \qquad \widehat{S}_1 = -\frac{i\beta\widehat{\mathcal{O}}}{2m_0c^2} = -\frac{i\beta\boldsymbol{\alpha}\cdot\widehat{\boldsymbol{\pi}}}{2m_0c}. \tag{B.7}$$

Then, the equation for $\Psi^{(1)}$ is

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Psi^{(1)} &= i\hbar\frac{\partial}{\partial t}(e^{i\widehat{S}_1}\Psi) = i\hbar\frac{\partial}{\partial t}(e^{i\widehat{S}_1})\Psi + e^{i\widehat{S}_1}(i\hbar\frac{\partial}{\partial t}\Psi) \\ &= \left[i\hbar\frac{\partial}{\partial t}(e^{i\widehat{S}_1}) + e^{i\widehat{S}_1}\widehat{H}_{\mathrm{D}}\right]\Psi \\ &= \left[i\hbar\frac{\partial}{\partial t}(e^{i\widehat{S}_1})e^{-i\widehat{S}_1} + e^{i\widehat{S}_1}\widehat{H}_{\mathrm{D}}e^{-i\widehat{S}_1}\right]\Psi^{(1)} \\ &= \left[e^{i\widehat{S}_1}\widehat{H}_{\mathrm{D}}e^{-i\widehat{S}_1} - i\hbar e^{i\widehat{S}_1}\frac{\partial}{\partial t}(e^{-i\widehat{S}_1})\right]\Psi^{(1)} \\ &= \widehat{H}_{\mathrm{D}}^{(1)}\Psi^{(1)} \end{aligned} \tag{B.8}$$

where we have used the identity $\frac{\partial}{\partial t}(e^{\widehat{A}})e^{-\widehat{A}} + e^{\widehat{A}}\frac{\partial}{\partial t}(e^{-\widehat{A}}) = \frac{\partial}{\partial t}\widehat{I} = 0$.

Now, using the identities

$$\begin{aligned} e^{\widehat{A}}\widehat{B}e^{-\widehat{A}} &= \widehat{B} + [\widehat{A}, \widehat{B}] + \frac{1}{2!}[\widehat{A}, [\widehat{A}, \widehat{B}]] + \frac{1}{3!}[\widehat{A}, [\widehat{A}, [\widehat{A}, \widehat{B}]]] \\ &\quad + \cdots e^{\widehat{A}(t)}\frac{\partial}{\partial t}(e^{-\widehat{A}(t)}) \\ &= \left(1 + \widehat{A}(t) + \frac{1}{2!}\widehat{A}(t)^2 + \frac{1}{3!}\widehat{A}(t)^3 \cdots\right) \\ &\quad \times\frac{\partial}{\partial t}\left(1 - \widehat{A}(t) + \frac{1}{2!}\widehat{A}(t)^2 - \frac{1}{3!}\widehat{A}(t)^3 \cdots\right) \\ &= \left(1 + \widehat{A}(t) + \frac{1}{2!}\widehat{A}(t)^2 + \frac{1}{3!}\widehat{A}(t)^3 \cdots\right) \\ &\qquad \times\left(-\frac{\partial\widehat{A}(t)}{\partial t} + \frac{1}{2!}\left\{\frac{\partial\widehat{A}(t)}{\partial t}\widehat{A}(t) + \widehat{A}(t)\frac{\partial\widehat{A}(t)}{\partial t}\right\}\right. \end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{3!}\left\{\frac{\partial \widehat{A}(t)}{\partial t}\widehat{A}(t)^2 + \widehat{A}(t)\frac{\partial \widehat{A}(t)}{\partial t}\widehat{A}(t) + \widehat{A}(t)^2\frac{\partial \widehat{A}(t)}{\partial t}\right\}\cdots\Bigg) \\
&\approx -\frac{\partial \widehat{A}(t)}{\partial t} - \frac{1}{2!}\left[\widehat{A}(t), \frac{\partial \widehat{A}(t)}{\partial t}\right] \\
&-\frac{1}{3!}\left[\widehat{A}(t), \left[\widehat{A}(t), \frac{\partial \widehat{A}(t)}{\partial t}\right]\right] \\
&-\frac{1}{4!}\left[\widehat{A}(t), \left[\widehat{A}(t), \left[\widehat{A}(t), \frac{\partial \widehat{A}(t)}{\partial t}\right]\right]\right], \qquad \text{(B.9)}
\end{aligned}
$$

with $\widehat{A} = i\widehat{S}_1$, we find

$$
\begin{aligned}
\widehat{H}_{\mathrm{D}}^{(1)} \approx\; & \widehat{H}_{\mathrm{D}} - \hbar\frac{\partial \widehat{S}_1}{\partial t} + i\left[\widehat{S}_1, \widehat{H}_{\mathrm{D}} - \frac{\hbar}{2}\frac{\partial \widehat{S}_1}{\partial t}\right] \\
& -\frac{1}{2!}\left[\widehat{S}_1, \left[\widehat{S}_1, \widehat{H}_{\mathrm{D}} - \frac{\hbar}{3}\frac{\partial \widehat{S}_1}{\partial t}\right]\right] \\
& -\frac{i}{3!}\left[\widehat{S}_1, \left[\widehat{S}_1, \left[\widehat{S}_1, \widehat{H}_{\mathrm{D}} - \frac{\hbar}{4}\frac{\partial \widehat{S}_1}{\partial t}\right]\right]\right]. \qquad \text{(B.10)}
\end{aligned}
$$

Substituting in (B.10), $\widehat{H}_{\mathrm{D}} = m_0c^2\beta + \widehat{\mathcal{E}} + \widehat{\mathcal{O}}$, simplifying the right-hand side using the relations $\beta\widehat{\mathcal{O}} = -\widehat{\mathcal{O}}\beta$ and $\beta\widehat{\mathcal{E}} = \widehat{\mathcal{E}}\beta$ and collecting everything together, we have

$$
\begin{aligned}
\widehat{H}_{\mathrm{D}}^{(1)} &\approx m_0c^2\beta + \widehat{\mathcal{E}}_1 + \widehat{\mathcal{O}}_1 \\
\widehat{\mathcal{E}}_1 &\approx \widehat{\mathcal{E}} + \frac{1}{2m_0c^2}\beta\widehat{\mathcal{O}}^2 - \frac{1}{8m_0^2c^4}\left[\widehat{\mathcal{O}}, \left([\widehat{\mathcal{O}}, \widehat{\mathcal{E}}] + i\hbar\frac{\partial \widehat{\mathcal{O}}}{\partial t}\right)\right] - \frac{1}{8m_0^3c^6}\beta\widehat{\mathcal{O}}^4 \\
\widehat{\mathcal{O}}_1 &\approx \frac{\beta}{2m_0c^2}\left([\widehat{\mathcal{O}}, \widehat{\mathcal{E}}] + i\hbar\frac{\partial \widehat{\mathcal{O}}}{\partial t}\right) - \frac{1}{3m_0^2c^4}\widehat{\mathcal{O}}^3, \qquad \text{(B.11)}
\end{aligned}
$$

with $\widehat{\mathcal{E}}_1$ and $\widehat{\mathcal{O}}_1$ obeying the relations $\beta\widehat{\mathcal{O}}_1 = -\widehat{\mathcal{O}}_1\beta$ and $\beta\widehat{\mathcal{E}}_1 = \widehat{\mathcal{E}}_1\beta$ exactly like $\widehat{\mathcal{E}}$ and $\widehat{\mathcal{O}}$. It is seen that while the term $\widehat{\mathcal{O}}$ in $\widehat{H}_{\mathrm{D}}$ is of order zero with respect to the expansion parameter $1/m_0c^2$ (*i.e.* $\widehat{\mathcal{O}} = O((1/m_0c^2)^0)$), the odd part of $\widehat{H}_{\mathrm{D}}^{(1)}$, namely $\widehat{\mathcal{O}}_1$, contains only terms of order $1/m_0c^2$ and higher powers of $1/m_0c^2$ (*i.e.*, $\widehat{\mathcal{O}}_1 = O((1/m_0c^2)))$.

To reduce the strength of the odd terms further in the transformed Hamiltonian a second Foldy–Wouthuysen transformation is applied with the same prescription:

$$
\Psi^{(2)} = e^{i\widehat{S}_2}\Psi^{(1)},
$$

$$
\begin{aligned}
\widehat{S}_2 &= -\frac{i\beta\widehat{\mathcal{O}}_1}{2m_0c^2} \\
&= -\frac{i\beta}{2m_0c^2}\left[\frac{\beta}{2m_0c^2}\left([\widehat{\mathcal{O}},\widehat{\mathcal{E}}] + i\hbar\frac{\partial\widehat{\mathcal{O}}}{\partial t}\right) - \frac{1}{3m_0^2c^4}\widehat{\mathcal{O}}^3\right].
\end{aligned} \tag{B.12}
$$

After this transformation,

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\Psi^{(2)} &= \widehat{H}_{\mathrm{D}}^{(2)}\Psi^{(2)}, \qquad \widehat{H}_{\mathrm{D}}^{(2)} = m_0c^2\beta + \widehat{\mathcal{E}}_2 + \widehat{\mathcal{O}}_2 \\
\widehat{\mathcal{E}}_2 &\approx \widehat{\mathcal{E}}_1, \qquad \widehat{\mathcal{O}}_2 \approx \frac{\beta}{2m_0c^2}\left(\left[\widehat{\mathcal{O}}_1,\widehat{\mathcal{E}}_1\right] + i\hbar\frac{\partial\widehat{\mathcal{O}}_1}{\partial t}\right),
\end{aligned} \tag{B.13}
$$

where now, $\widehat{\mathcal{O}}_2 = O((1/m_0c^2)^2)$. After the third transformation

$$
\Psi^{(3)} = e^{i\widehat{S}_3}\,\Psi^{(2)}, \qquad \widehat{S}_3 = -\frac{i\beta\widehat{\mathcal{O}}_2}{2m_0c^2}, \tag{B.14}
$$

we have

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\Psi^{(3)} &= \widehat{H}_{\mathrm{D}}^{(3)}\Psi^{(3)}, \qquad \widehat{H}_{\mathrm{D}}^{(3)} = m_0c^2\beta + \widehat{\mathcal{E}}_3 + \widehat{\mathcal{O}}_3 \\
\widehat{\mathcal{E}}_3 &\approx \widehat{\mathcal{E}}_2 \approx \widehat{\mathcal{E}}_1, \qquad \widehat{\mathcal{O}}_3 \approx \frac{\beta}{2m_0c^2}\left([\widehat{\mathcal{O}}_2,\widehat{\mathcal{E}}_2] + i\hbar\frac{\partial\widehat{\mathcal{O}}_2}{\partial t}\right),
\end{aligned} \tag{B.15}
$$

where $\widehat{\mathcal{O}}_3 = O((1/m_0c^2)^3)$. So, neglecting $\widehat{\mathcal{O}}_3$,

$$
\begin{aligned}
\widehat{H}_{\mathrm{D}}^{(3)} \approx{}& m_0c^2\beta + \widehat{\mathcal{E}} + \frac{1}{2m_0c^2}\beta\widehat{\mathcal{O}}^2 \\
&- \frac{1}{8m_0^2c^4}\left[\widehat{\mathcal{O}},\left([\widehat{\mathcal{O}},\widehat{\mathcal{E}}] + i\hbar\frac{\partial\widehat{\mathcal{O}}}{\partial t}\right)\right] \\
&- \frac{1}{8m_0^3c^6}\beta\left\{\widehat{\mathcal{O}}^4 + \left([\widehat{\mathcal{O}},\widehat{\mathcal{E}}] + i\hbar\frac{\partial\widehat{\mathcal{O}}}{\partial t}\right)^2\right\}
\end{aligned} \tag{B.16}
$$

It may be noted that starting with the second transformation successive $(\widehat{\mathcal{E}}, \widehat{\mathcal{O}})$ pairs can be obtained recursively using the rule

$$
\begin{aligned}
\widehat{\mathcal{E}}_j &= \widehat{\mathcal{E}}_1\left(\widehat{\mathcal{E}} \to \widehat{\mathcal{E}}_{j-1}, \widehat{\mathcal{O}} \to \widehat{\mathcal{O}}_{j-1}\right) \\
\widehat{\mathcal{O}}_j &= \widehat{\mathcal{O}}_1\left(\widehat{\mathcal{E}} \to \widehat{\mathcal{E}}_{j-1}, \widehat{\mathcal{O}} \to \widehat{\mathcal{O}}_{j-1}\right), \quad j > 1,
\end{aligned} \tag{B.17}
$$

and retaining only the relevant terms of desired order at each step.

With $\widehat{\mathcal{E}} = q\phi$ and $\widehat{\mathcal{O}} = c\boldsymbol{\alpha} \cdot \widehat{\boldsymbol{\pi}}$, the final reduced Hamiltonian (B.16) is, to the order calculated,

$$\widehat{H}_{\mathrm{D}}^{(3)} = \beta \left(m_0c^2 + \frac{\widehat{\pi}^2}{2m_0} - \frac{\widehat{p}^4}{8m_0^3c^6} \right) + q\phi - \frac{q\hbar}{2m_0c}\beta \boldsymbol{\Sigma} \cdot \mathbf{B}$$

$$- \frac{iq\hbar^2}{8m_0^2c^2}\boldsymbol{\Sigma} \cdot \mathrm{curl}\, \mathbf{E} - \frac{q\hbar}{4m_0^2c^2}\boldsymbol{\Sigma} \cdot \mathbf{E} \times \widehat{\mathbf{p}} - \frac{q\hbar^2}{8m_0^2c^2}\mathrm{div}\, \mathbf{E}, \quad \text{(B.18)}$$

with the individual terms having direct physical interpretations. The terms in the first parenthesis result from the expansion of $\sqrt{m_0^2c^4 + c^2\widehat{\pi}^2}$ showing the effect of the relativistic mass increase. The second and third terms are the electrostatic and magnetic dipole energies. The next two terms, taken together (for hermiticity), contain the spin–orbit interaction. The last term, the so-called Darwin term, is attributed to the *zitterbewegung* (trembling motion) of the Dirac particle: because of the rapid coordinate fluctuations over distances of the order of the Compton wavelength $(2\pi\hbar/m_0c)$, the particle sees a somewhat smeared out electric potential.

It is clear that the Foldy–Wouthuysen transformation technique expands the Dirac Hamiltonian as a power series in the parameter $1/m_0c^2$, enabling the use of a systematic approximation procedure for studying the deviations from the nonrelativistic situation. We note the analogy between the nonrelativistic particle dynamics and paraxial optics:

Standard Dirac equation	Beam-optical form
$m_0c^2\beta + \widehat{\mathcal{E}}_{\mathrm{D}} + \widehat{\mathcal{O}}_{\mathrm{D}}$	$-n_0\sigma_z + \widehat{\mathcal{E}} + \widehat{\mathcal{O}}$
m_0c^2	$-n_0$
Positive energy	Forward propagation
Nonrelativistic, $\lvert\widehat{\boldsymbol{\pi}}\rvert \ll m_0c$	Paraxial beam, $\lvert\widehat{\mathbf{p}}_\perp\rvert \ll n_0$
Nonrelativistic motion + relativistic corrections	Paraxial behavior + aberration corrections

Noting the earlier analogy, the idea of Foldy–Wouthuysen form of the Dirac theory has been adopted to study the paraxial optics and deviations from it by first casting the Maxwell equations in a spinor form resembling exactly the Dirac equations (B.1) and (B.2) in all respects: *i.e.*, a multicomponent Ψ having the upper half of its components large compared to the lower components and the Hamiltonian having an even part $(\widehat{\mathcal{E}})$, an odd part $(\widehat{\mathcal{O}})$, a suitable expansion parameter, $(\lvert\widehat{\mathbf{p}}_\perp\rvert/n_0 \ll 1)$ characterizing the dominant forward propagation and a leading term with a β coefficient (σ_z in our formalism) commuting with $\widehat{\mathcal{E}}$ and anticommuting with $\widehat{\mathcal{O}}$. The additional feature of our formalism is to return finally to the original representation after making an extra approximation, dropping σ_z

from the final reduced optical Hamiltonian, taking into account the fact that we are primarily interested only in the forward-propagating beam.

APPENDIX C: THE MAGNUS FORMULA

The Magnus formula is the continuous analogue of the famous Baker–Campbell–Hausdorff (BCH) formula

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}+\frac{1}{2}[\hat{A},\hat{B}]+\frac{1}{12}\{[[\hat{A},\hat{A}],\hat{B}]+[[\hat{A},\hat{B}],\hat{B}]\}+\cdots}. \tag{C.1}$$

Let it be required to solve the differential equation

$$\frac{\partial}{\partial t}u(t) = \hat{A}(t)u(t) \tag{C.2}$$

to get $u(T)$ at $T > t_0$, given the value of $u(t_0)$; the operator $\hat{A}$ can represent any linear operation. For an infinitesimal Δt, we can write

$$u(t_0 + \Delta t) = e^{\Delta t \hat{A}(t_0)}u(t_0). \tag{C.3}$$

Iterating this solution we have

$$\begin{aligned} u(t_0 + 2\Delta t) &= e^{\Delta t \hat{A}(t_0+\Delta t)}e^{\Delta t \hat{A}(t_0)}u(t_0) \\ u(t_0 + 3\Delta t) &= e^{\Delta t \hat{A}(t_0+2\Delta t)}e^{\Delta t \hat{A}(t_0+\Delta t)}e^{\Delta t \hat{A}(t_0)}u(t_0)\cdots \quad \text{and so on.} \end{aligned} \tag{C.4}$$

If $T = t_0 + N\Delta t$ we would have

$$u(T) = \left\{\prod_{n=0}^{N-1} e^{\Delta t \hat{A}(t_0+n\Delta t)}\right\} u(t_0). \tag{C.5}$$

Thus, $u(T)$ is given by computing the product in (C.5) using successively the BCH formula (C.1) and considering the limit $\Delta t \longrightarrow 0$, $N \longrightarrow \infty$ such that $N\Delta t = T - t_0$. The resulting expression is the Magnus formula (Magnus, 1954):

$$\begin{aligned} u(T) &= \widehat{\mathcal{T}}(T, t_0)u(t_0) \\ \mathcal{T}(T, t_0) &= \exp\left\{\int_{t_0}^{T} dt_1\, \hat{A}(t_1)\right. \\ &\quad + \frac{1}{2}\int_{t_0}^{T} dt_2 \int_{t_0}^{t_2} dt_1\, [\widehat{A}(t_2), \hat{A}(t_1)] \\ &\quad + \frac{1}{6}\int_{t_0}^{T} dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1\, ([[\hat{A}(t_3), \hat{A}(t_2)], \hat{A}(t_1)] \\ &\quad \left. + [[\hat{A}(t_1), \hat{A}(t_2)], \hat{A}(t_3)]) + \cdots \right\}. \end{aligned} \tag{C.6}$$

To see how the Eq. (C.6) is obtained, let us substitute the assumed form of the solution, $u(t) = \widehat{\mathcal{T}}\,(t, t_0)\,u\,(t_0)$, in (C.2). Then, it is seen that $\widehat{\mathcal{T}}(t, t_0)$ obeys the equation

$$\frac{\partial}{\partial t}\widehat{\mathcal{T}}(t, t_0) = \widehat{A}(t)\mathcal{T}(t, t_0), \qquad \widehat{\mathcal{T}}(t_0, t_0) = \widehat{I}. \tag{C.7}$$

Introducing an iteration parameter λ in (C.7), let

$$\frac{\partial}{\partial t}\widehat{\mathcal{T}}(t, t_0; \lambda) = \lambda\widehat{A}(t)\widehat{\mathcal{T}}(t, t_0; \lambda), \tag{C.8}$$

$$\widehat{\mathcal{T}}(t_0, t_0; \lambda) = \widehat{I}, \quad \widehat{\mathcal{T}}(t, t_0; 1) = \widehat{\mathcal{T}}(t, t_0). \tag{C.9}$$

Assume a solution of (C.8) to be of the form

$$\widehat{\mathcal{T}}(t, t_0; \lambda) = e^{\Omega(t, t_0; \lambda)} \tag{C.10}$$

with

$$\Omega(t, t_0; \lambda) = \sum_{n=1}^{\infty} \lambda^n \Delta_n(t, t_0), \quad \Delta_n(t_0, t_0) = 0 \quad \text{for all } n. \tag{C.11}$$

Now, using the identity (Wilcox, 1967)

$$\frac{\partial}{\partial t}e^{\Omega(t,t_0;\lambda)} = \left\{\int_0^1 ds e^{s\Omega(t,t_0;\lambda)}\frac{\partial}{\partial t}\Omega(t, t_0; \lambda)e^{-s\Omega(t,t_0;\lambda)}\right\} e^{\Omega(t,\lambda)}, \tag{C.12}$$

one has

$$\int_0^1 ds e^{s\Omega(t,t_0;\lambda)}\frac{\partial}{\partial t}\Omega(t, t_0; \lambda)e^{-s\Omega(t,t_0;\lambda)} = \lambda\widehat{A}(t). \tag{C.13}$$

Substituting in (C.13) the series expression for $\Omega(t, t_0; \lambda)$ (C.11), expanding the left-hand side using the first identity in (C8), integrating and equating the co-efficients of λ^j on both sides, we get, recursively, the equations for $\Delta_1(t, t_0)$, $\Delta_2(t, t_0), \ldots$, etc. For $j = 1$

$$\frac{\partial}{\partial t}\Delta_1(t, t_0) = \hat{A}(t), \qquad \Delta_1(t_0, t_0) = 0 \tag{C.14}$$

and hence

$$\Delta_1(t, t_0) = \int_{t_0}^{t} dt_1 \widehat{A}(t_1). \tag{C.15}$$

For $j = 2$

$$\frac{\partial}{\partial t}\Delta_2(t, t_0) + \frac{1}{2}\left[\Delta_1(t, t_0), \frac{\partial}{\partial t}\Delta_1(t, t_0)\right] = 0, \qquad \Delta_2(t_0, t_0) = 0 \tag{C.16}$$

and hence

$$\Delta_2(t, t_0) = \frac{1}{2}\int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1 [\widehat{A}(t_2), \widehat{A}(t_1)]. \tag{C.17}$$

Similarly,

$$\Delta_3(t, t_0) = \frac{1}{6}\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 \{[[\hat{A}(t_1),\ \hat{A}(t_2)], \hat{A}(t_3)] + [[\hat{A}(t_3),\ \hat{A}(t_2)], \hat{A}(t_1)]\}. \tag{C.18}$$

Then, the Magnus formula in (C.6) follows from (C.9)–(C.11). Equation (32) we have, in the context of z-evolution follows from the earlier discussion with the identification $t \longrightarrow z$, $t_0 \longrightarrow z^{(1)}$, $T \longrightarrow z^{(2)}$ and $\hat{A}(t) \longrightarrow -\frac{i}{\hbar}\widehat{\mathcal{H}}_o(z)$.

For more details on the exponential solutions of linear differential equations, related operator techniques and applications to physical problems the reader is referred to Wilcox (1967), Bellman and Vasudevan (1986), Dattoli *et al.* (1993), and references therein.

APPENDIX D: ANALOGIES BETWEEN LIGHT OPTICS AND CHARGED-PARTICLE OPTICS: RECENT DEVELOPMENTS

Historically, variational principles have played a fundamental role in the evolution of mathematical models in classical physics, and many equations can be derived by using them. Here the relevant examples are Fermat's principle in optics and Maupertuis' principle in mechanics. The beginning of the analogy between geometrical optics and mechanics is usually attributed to Descartes (1637), but actually it can traced back to Ibn Al-Haitham Alhazen (965–1037) (Ambrosini *et al.*, 1997). The analogy between the trajectory of material particles in potential fields and the path of light rays in media with continuously variable refractive index was formalized by Hamilton in 1833. The Hamiltonian analogy lead to the development of electron optics in 1920s, when Busch derived the focusing action and a lens-like action of the axially symmetric magnetic field using the methodology of geometrical optics. Around the same time, Louis de Broglie associated his now famous wavelength to moving particles. Schrödinger extended the analogy by passing from geometrical optics to wave optics through his wave equation incorporating the de Broglie wavelength. This analogy played a fundamental role in the early development of quantum mechanics. The analogy, on the other hand, lead to the development of practical electron optics and one of the early inventions was the electron microscope by Ernst Ruska. A detailed account of Hamilton's analogy is available in (Born and Wolf, 1999; Hawkes and Kasper, 1989).

Until very recently, it was possible to see this analogy only between the geometrical-optic and classical prescriptions of electron optics. The reasons being that, the quantum theories of charged-particle beam optics have been under development only for about a decade (Jagannathan *et al.*, 1989; Jagannathan, 1990, 1993; Khan and Jagannathan, 1995; Jagannathan and Khan, 1996, 1997, 1998; Conte *et al.*, 1996; Khan, 1997) with the very expected feature of wavelength-dependent effects, which have no analogue in the traditional descriptions of light beam optics. With the current development of the non-traditional prescriptions of Helmholtz optics (Khan *et al.*, 2002) and the matrix formulation of Maxwell optics, accompanied with wavelength-dependent effects, it is seen that the analogy between the two systems persists. The non-traditional prescription of Helmholtz optics is in close analogy with the quantum theory of charged-particle beam optics based on the Klein–Gordon equation. The matrix formulation of Maxwell optics is in close analogy with the quantum theory of charged-particle beam optics based on the Dirac Equation (Khan, 2002).

REFERENCES

Acharya, R. and Sudarshan, E. C. G. (1960). Front description in relative quantum mechanics, *Journal of Mathematical Physics* **1**, 532–536.

Ambrosini, D., Ponticiello, A., Schirripa Spagnolo, G., Borghi, R., and Gori, F. (1997). Bouncing light beams and the Hamiltonian analogy. *European Journal of Physics* **18**, 284–289.

Bellman, R. and Vasudevan, R. (1986). *Wave Propagation: An Invariant Imbedding Approach*, Reidel, Dordrecht.

Bjorken, J. D. and Drell, S. D. (1964). *Relativistic Quantum Mechanics*, McGraw-Hill, New York, San Francisco.

Born, M. and Wolf, E. (1999). *Principles of Optics*, 7th edn., Cambridge University Press, United Kingdom.

Chen, P. (1999) (ed.). *Proceedings of the 15th Advanced ICFA Beam Dynamics Workshop on Quantum Aspects of Beam Physics*, 4–9 January 1998, Monterrey, California, USA, World Scientific, Singapore, http://www.slac.stanford.edu/grp/ara/qabp/qabp.html; Chen, P. (2002) (ed.), *Proceedings of the 18th Advanced ICFA Beam Dynamics Workshop on Quantum Aspects of Beam Physics*, 15–20 October 2000, Capri, Italy, World Scientific, Singapore, http://qabp2k.sa.infn.it/; Chen, P. (2003) (ed.). *Proceedings of the Joint 28th ICFA Advanced Beam Dynamics and Advanced & Novel Accelerators Workshop on QUANTUM ASPECTS OF BEAM PHYSICS and Other Critical Issues of Beams in Physics and Astrophysics*, 7–11 January 2003, Hiroshima University, Japan, World Scientific, Singapore, http://home.hiroshima-u.ac.jp/ogata/qabp/home.html; Workshop Reports: *ICFA Beam Dynamics Newsletter* **16**, (1988) 22–25; *ibid* **23**, (2000) 13–14; *ibid* **30**, (2003) 72–75; *Bulletin of the Association of Asia Pacific Physical Societies* **13**(1), (2003) 34–37.

Conte, M., Jagannathan, R., Khan, S. A., and Pusterla, M. (1996). Beam optics of the Dirac particle with anomalous magnetic moment. *Particle Accel.* **56**, 99–126.

Dattoli, G., Renieri, A., and Torre, A. (1993). *Lectures on the Free Electron Laser Theory and Related Topics*, World Scientific, Singapore.

Dragt, A. J., Forest, E., and Wolf, K. B. (1986). *Lie Methods in Optics*, *Lecture Notes in Physics*, Vol. 250, Springer Verlag, pp. 105–157.

Dragt, A. J. and Forest, E. (1986). *Adv. Imag. Electron Phys.* **67**, 65–120; Dragt, A. J., Neri, F., Rangarajan, G., Douglas, D. R., Healy, L. M., and Ryne, R. D. (1988). Lie algebriac treatment of linear and nonlinear beam dynamics. *Ann. Rev. Nucl. Part. Sci.* **38**, 455–496; Forest, E. and Hirata, K. (1992). *A Contemporary Guide to Beam Dynamics*, KEK Report 92-12, National Laboratory for High Energy Physics, Tsukuba, Japan; Forest, E., Berz, M., and Irwin, J. (1989). *Particle Accel.* **24**, 91–97; Rangarajan, G., Dragt, A. J., and Neri, F. (1990). Solvable map representation of a nonlinear sympletic map. *Particle Accel.* **28**, 119–124; Ryne, R. D. and Dragt, A. J. (1991). Magnetic optics calculations for cylinderically symmetric beams. *Particle Accel.* **35**, 129–165.

Dragt, A. J. (1998). Lie algebriac theory of geometrical optics and optical aberrations. *J. Opt. Soc. Am* **72**, (1982) 372; *Lie Algebraic Method for Ray and Wave Optics*, University of Maryland Physics Department Report.

Fedele, R. and Man'ko, V. I. (1999). The role of semiclassical description in the quantum-like theory of light rays. *Physical Review E* **60**, 6042–6050.

Feshbach, H. and Villars, F. M. H. (1958). Elementary relativistic wave mechanics of spin 0 and spin 1/2 particles. *Reviews of Modern Physics* **30**, 24–45.

Fishman, L. and McCoy, J. J. (1984). Derivation and application of extended parabolic wave theories. Part I. The factored Helmholtz equation. *Journal of Mathematical Physics* **25**, 285–296.

Foldy, L. L. and Wouthuysen, S. A. (1950). On the Dirac theory of spin 1/2 particles and its non-relativistic limit. *Physical Review* **78**, 29–36.

Goodman, J. W. (1996). *Introduction to Fourier Optics*, 2nd edn., McGraw-Hill, New York.

Hawkes, P. W. and Kasper, E. (1989). *Principles of Electron Optics*, Vols. I and II, Academic Press, London; Hawkes, P. W. and Kasper, E. (1994) *Principles of Electron Optics*. Vol. 3: *Wave Optics*, Academic Press, London and San Diego.

Jagannathan, R., Simon, R., Sudarshan, E. C. G., and Mukunda, N. (1989). Quantum theory of magnetic electron lenses based on the Dirac equation. *Physics Letters A* **134**, 457–464.

Jagannathan, R. (1990). Quantum theory of electron lenses based on the Dirac equation. *Physics Letters A* **42**, 6674–6689.

Jagannathan, R. (1993). Dirac equation and electron optics. In Dutt, R. and Ray, A. K. (eds.), *Dirac and Feynman: Pioneers in Quantum Mechanics*, Wiley Eastern, New Delhi, India, pp. 75–82.

Jagannathan, R. and Khan, S. A. (1996). Quantum theory of the optics of charged particles. In Hawkes Peter, W. (ed.), *Advances in Imaging and Electron Physics*, Vol. 97, Academic Press, San Diego, pp. 257–358.

Jagannathan, R. and Khan, S. A. (1997). Quantum mechanics of accelerator optics. *ICFA Beam Dynamics Newsletter* **13**, 21–27 (ICFA: International Committee for Future Accelerators).

Jagannathan, R. and Khan, S. A. (1998). Several articles in Proceedings/E-Prints on the Quantum theory of charged-particle beam optics, *arXiv: physics/9803042*; *arXiv: physics/0101060*; *arXiv: physics/9809032*; *arXiv: physics/9904063*; *arXiv: physics/0112085*; *arXiv: physics/0112086* and *arXiv: physics/0304099*.

Khan, S. A. and Jagannathan, R. (1995). On the quantum mechanics of charged particle beam transport through magnetic lenses. *Physical Review E* **51**, 2510–2515.

Khan, S. A. (1997). *Quantum Theory of Charged-Particle Beam Optics*, Ph.D. Thesis, University of Madras, Chennai, India.

Khan, S. A., Jagannathan, R., and Simon, R. (2002). Foldy-Wouthuysen transformation and a quasi-paraxial approximation scheme for the scalar wave theory of light beams, *arXiv: physics/0209082* (*communicated*).

Khan, S. A. (2002). Analogies between light optics and charged-particle optics. *ICFA Beam Dynamics Newsletter* **27**, 42–48; *arXiv: physics/0210028* (ICFA: International Committee for Future Accelerators).

Magnus, W. (1954). On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics* **7**, 649–673.

Pryce, M. H. L. (1948). The mass-centre in the restricted theory of relativity and its connexion with the quantum theory of elementary particles. *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences* **195**, 62–81.

Tani, S. (1951). Connection between particle models and field theories. I. The case spin 1/2. *Progress of Theoretical Physics* **6**, 267–285.

Todesco, E. (1999). Overview of single-particle nonlinear dynamics, CERN-LHC-99-1-MMS, 16pp; Talk given at 16th ICFA Beam Dynamics Workshop on Nonlinear and Collective Phenomena in Beam Physics, Arcidosso, Italy, 1–5 September 1998; *AIP Conference Proceedings* **468**, 157–172.

Turchetti, G., Bazzani, A., Giovannozzi, M., Servizi, G., and Todesco, E. (1989). Normal forms for symplectic maps and stability of beams in particle accelarators. In *Proceedings of the Dynamical symmetries and Chaotic Behaviour in Physical Systems*, Bologna, Italy, pp. 203–231.

Wilcox, R. M. (1967). Exponential operators and parameter differentiation in quantum physics. *Journal of Mathematical Physics* **8**(4), 962–982.