Quantum theory of electron lenses based on the Dirac equation

R. Jagannathan
The Institute of Mathematical Sciences, Central Institutes of Technology Campus, Tharamani P.O., Madras 600113, India
(Received 19 March 1990; revised manuscript received 21 June 1990)

The quantum theory of a magnetic electron lens with rotational symmetry about its straight optic axis has recently been studied entirely on the basis of the Dirac equation [R. Jagannathan, R. Simon, E. C. G. Sudarshan, and N. Mukunda, Phys. Lett. A **134**, 457 (1989)]. Following the same type of algebraic approach, the present paper elaborates on the quantum mechanics of beam transport through a general electron-lens system with a straight axis. The basic theory is developed in a general framework, and examples of applications are treated under the paraxial approximation and thin-lens assumption. Extension of the theory to systems with curved optic axes is dealt with very briefly. The present formalism based on a "beam-optical" representation of the Dirac equation goes over directly to the Hamiltonian formalism of geometrical electron optics upon classicalization.

I. INTRODUCTION

The quantum mechanics of a magnetic electron lens with rotational symmetry about its straight optic axis has recently been studied entirely on the basis of the Dirac equation.[1] The present paper elaborates on the quantum mechanics of beam transport through a general electron-lens system with straight axis following the same type of algebraic approach as in Ref. 1. The basic theory is developed in a general framework and examples of applications (axially symmetric magnetic lens, magnetic quadrupole lens, axially symmetric electrostatic lens, electrostatic quadrupole lens) are treated under the paraxial approximation and thin-lens assumption. Extension of the basic theory to systems with curved optic axes is treated very briefly.

In studying the behavior of any electron optics system major interest is in the evolution of the beam parameters along the system axis. Hence, in the spinor formalism, the most convenient representation of the Dirac equation to work with would be one that specifies directly, and as simply as possible, the evolution of the wave function along the axis. It is found that in the case of propagation of a monoenergetic electron beam the corresponding time-independent Dirac equation can be cast in the desired form by a simple transformation. The resulting "beam-optical" representation of the Dirac equation forms the basis for our treatment of electron optics at the level of single-particle dynamics (initiated in Ref. 1 and developed in the present work). Starting with this representation of the Dirac equation, and using well-known algebraic techniques, we obtain a formalism which allows one to study the given electron optics system successively under the paraxial and higher-order approximations. This approach leads directly upon "classicalization" to the familiar Hamiltonian formalism of geometrical electron optics (relativistic, including aberrations). Further, it may be noted that the basic framework of our theory is constructed, *ab initio*, starting from the Dirac equation and entirely based on it; of course, the electromagnetic fields of the system are to be described in classical terms by the same expressions, and approximations, as in classical electron optics. The only "basic approximation" that underlies our formalism arises from the assumption that we are dealing with a positive-energy quasiparaxial electron beam propagating along the optic axis of the system always "forward directed."

To place our work in proper perspective, some remarks are in order. In the context of electron microscopy, the conventional description of the quantum mechanics of image formation stems from the pioneering work of Glaser[2] based on a semiclassical treatment of the nonrelativistic Schrödinger equation and modeled after the traditional Fresnel-Kirchhoff integral approach of light optics. (For excellent reviews of the electron optical transfer theory presently used, at introductory and advanced levels, cf. Ref. 3, wherein the details of the development of the theory and extensive bibliography on the topic are also available.) To deal with systems employing electron beams at relativistic energies, the normal suggestion[4] is just to use the nonrelativistic scalar description with substitutions for the mass and the de Broglie wavelength of the electron by "relativistically corrected" expressions:

$$m \rightarrow m(1-v^2/c^2)^{-1/2}, \quad \lambda \rightarrow \lambda(1-v^2/c^2)^{1/2} .$$

This suggestion comes mainly from some theoretical studies[5] on the relativistic treatment of the interaction between the beam electrons and the specimen being examined in the microscope, using the Dirac equation and the first Born approximation of the scattering theory, and takes no account of spin. That spin does play a role in the optics of electrons was clear from the preliminary work of Rubinowicz[6] and later studies of Phan Van Loc[7] on the diffraction of free Dirac electron waves. Rubinowicz studied the diffraction wave in the Kirchhoff-type theory of Dirac electron waves. Phan Van Loc contributed significantly to the wave optics of Dirac electrons by establishing the general form of the Huygens principle in

 ©1990 The American Physical Society

the case of free electrons, using a very elegant mathematical formalism, and applying it to the study of diffraction by apertures and other obstacles; this, incidentally, demonstrated the result of Rubinowicz for waves with harmonic time dependence [$\sim\exp(-i\omega t)$] by a different and simpler method. This discovery of the role of spin in diffraction does affect the conclusions of Glaser[2] which are based on the Schrödinger equation. However, after the work of Phan Van Loc, for a long time, there seems to have been no further serious consideration of the Dirac equation in connection with electron optics; perhaps, the need for such a study was not strongly felt then, from the point of view of practical applications. But, with the present-day electron microscopes commonly operating at accelerating voltages of the order of one hundred to several hundreds of kilovolts (even in *MV* region), it is only proper that relativistic effects are now considered seriously. So, there has arisen the need to reexamine the usual linear transfer theory of electron optics with a nonrelativistic basis, using the appropriate quantum-mechanical equation, namely, the Dirac equation. In fact, one may recall here that in practice, as per the past experience,[8] scaled-up versions of the standard 100-kV designs operating at much higher accelerating voltages fail to achieve a resolving power comparable to that of the best 100-keV designs. This situation has prompted the recent investigations of Ferwerda, Hoenders, and Slump,[9] addressing, for the first time, the problem of studying the propagation of the Dirac wave function through an electron microscope, in order to initiate the development of a relativistic quantum theory of electron optical image formation. To this end, they have considered it adequate to treat spin as a spectator degree of freedom and to use a semiclassical ($\hbar\rightarrow 0$) approach closely following Glaser's nonrelativistic theory. The arguments for this are as follows: (i) The electromagnetic fields comprising the optical elements of the electron microscopes are not usually strong enough to affect the spin of the electrons very much. (ii) The spin states of the electrons are not, anyway, to be recorded in the observation of the image. (iii) The spin-dependent contribution[6,7] to the propagation of the Dirac wave function vanishes in the limit $\hbar\rightarrow 0$. (iv) The electromagnetic potentials do not vary appreciably over distances of the order of a Compton wavelength of the electron so that the WKB approximation can be used. Then, eventually, we are led to the conclusion that to account for electron optical image formation in the relativistic case, a scalar theory based on the Klein-Gordon equation, treated using a semiclassical analysis exactly along the lines of Glaser's nonrelativistic theory, is fully equipped. Consequently, it is suggested that Glaser's nonrelativistic theory could be used with minor modifications for the treatment of the relativistic case: the classical nonrelativistic eikonal function in the kernal of Glaser's integral relation connecting the wave functions at two different planes along the optic axis of the system is to be now replaced by the eikonal function calculated using the solutions of relativistic trajectory equations. Thus, it is seen that the fundamental study of Ferwerda, Hoenders, and Slump[9] replaces the rather ad hoc substitution recipes customarily used for extending the nonrelativistic scalar quantum theory to the relativistic domain, by a systematic procedure well founded within the limitations of its semiclassical basis and the approximations involved. Such a semiclassical scalar theory, effectively ignoring the spinor character of the electron wave function (that too at relativistic energies), might be "sufficient" to get optimal designs for high-voltage microscopes but, as yet, it is not clear. However, certainly it should be considered worthwhile to have a formalism in which one does not regard the orbital and spin degrees of freedom as if they exist separately, or, in other words, the four-component spinor wave function is handled as a single entity obeying the Dirac equation; once there is such a formalism, as general as possible, approximations can always be used as are appropriate to the particular system under study. This is what our work aims to achieve. As already mentioned above, the only "basic approximation" underlying our formalism is derived from the assumption that we are dealing with a quasiparaxial positive-energy electron beam propagating along the system axis always "forward directed." Interestingly, it turns out, as will be seen below, that there is a nontrivial difference in the mathematical accounts of the action of a lens on the electron wave function depending on whether one is using the spinor theory or a scalar approximation thereof.

In the context of accelerator physics, while rigorous quantum theory is used for studying radiation, spin dynamics, and various quantum effects, the traditional treatment of the orbital motion, or the "optics," of the high-energy charged particle beams, at the level of single-particle dynamics, is usually based on classical physics.[10] As already mentioned above, the quantum-mechanical formalism of electron optics based on the Dirac equation at the level of single-particle dynamics, obtained in our work, goes over in the classical limit, naturally, to the Hamiltonian formalism of geometrical electron optics which forms the basis of the present-day accelerator optics. So, it may be interesting to see whether the quantum-mechanical formalism would furnish any new insight into the foundations of accelerator optics.

The plan of the paper is as follows. The general formalism of the basic theory of beam transport in electron lens systems with straight optic axis is developed in Sec. II entirely on the basis of the Dirac equation. Specific examples are taken up Sec. III. The method of extending the treatment to systems with curved optic axes is sketched briefly in Sec. IV. The paper concludes in Sec. V with several remarks on the passage from quantum theory to the classical theory of electron optics, scalar approximations of the formalism based on the Dirac theory leading to the Ferwerda-Hoenders-Slump (FHS) formalism in the relativistic case and Glaser's theory in the nonrelativistic case, and the analogy with light optics.

II. GENERAL THEORY OF BEAM TRANSPORT IN AN ELECTRON-LENS SYSTEM WITH STRAIGHT OPTIC AXIS

Let the electron-lens system under study be comprised of the static electromagnetic field $\{\mathbf{E}(\mathbf{x}),\mathbf{B}(\mathbf{x})\}$ specified by the four-vector potential $\{\Phi(\mathbf{x}), \mathbf{A}(\mathbf{x})\}$, and the system (optic) axis be taken as the z axis close to which the beam propagates. The system does not have, in principle, a sharp boundary. But, for all practical purposes, it may be considered to be situated on the optic axis between two transverse planes at, say, $z=z_l$ and $z=z_r$ $(z_l<z_r)$, i.e., the region outside the interval (z_l,z_r) is practically field free.

The four-component spinor wave function of the electron beam propagating through the system is subject to the Dirac equation

$$i\hbar\frac{\partial\Psi}{\partial t}(\mathbf{x},t)=(mc^2\beta-e\Phi+c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}})\Psi(\mathbf{x},t)=\hat{H}_D\Psi(\mathbf{x},t)\ , \tag{1}$$

$$\beta=\begin{bmatrix} I & 0 \\ 0 & -I\end{bmatrix},\quad \boldsymbol{\alpha}=\begin{bmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0\end{bmatrix},\quad \hat{\boldsymbol{\pi}}=\hat{\mathbf{p}}+(e/c)\mathbf{A}\ .$$

Let us take the input beam for the system (i.e., in the region $z<z_l$) to be almost paraxial (close to the axis), and quasimonoenergetic with p_0 as the mean momentum of the beam electrons. We assume that the system field is such that the output is again a "beam," with changes in the shape due to the lens action of the system. We are dealing with the scattering states of the system with a time-independent Hamiltonian. Hence the solutions of (1) relevant for us, representing almost paraxial quasimonoenergetic beams moving through the system along the $+z$ axis, are of the form

$$\begin{aligned}
&\Psi(\mathbf{x},t)=\int_{p_0-\Delta p}^{p_0+\Delta p} dp\, \exp[-(i/\hbar)E(p)t]\psi(\mathbf{x};p)\ ,\\
&\qquad\qquad p=|\mathbf{p}|,\quad \Delta p \ll p_0\ ,\\
&E(p)=+(m^2c^4+c^2p^2)^{1/2}\ ,\\
&\psi(\mathbf{x}_\perp,z;p)\xrightarrow{z\to z_l}\psi_{(\mathrm{in})}(\mathbf{x}_\perp,z;p)\\
&\qquad=\int d^2p_\perp\,\mathcal{U}(\mathbf{p})\exp[(i/\hbar)\mathbf{p}\cdot\mathbf{x}]\ ,\\
&\qquad\qquad |\mathbf{p}_\perp|\ll p,\quad \mathbf{p}=(\mathbf{p}_\perp,+(p^2-p_\perp^2)^{1/2})\\
&\mathcal{U}(\mathbf{p})=[a_+(\mathbf{p})u_+(\mathbf{p})+a_-(\mathbf{p})u_-(\mathbf{p})]\ ,
\end{aligned} \tag{2}$$

with

$$\{u_\pm(\mathbf{p})\exp[(i/\hbar)(\mathbf{p}\cdot\mathbf{x}-E(p)t)]\}$$

as the standard positive-energy plane-wave solutions of the free-electron Dirac equation.

Our aim is to understand the quantum mechanics of an electron-lens system with the z axis as its optic axis, i.e., our interest is in studying the z evolution of the electron beam wave function $\Psi(\mathbf{x},t)$ of the form (2) subject to the Dirac equation (1). In other words, we want to find the relation between $\Psi(\mathbf{x}_\perp,z',t)$ and $\Psi(\mathbf{x}_\perp,z'',t)$, respectively, the wave functions (optical) at the transverse (profile) planes at positions z' and z'' on the axis; particularly of interest are the profile planes at $z'<z_l$ (the input space) and at $z''>z_r$ (the output space). To this end, let us seek a relation of the type

$$\psi(z'';p)=\hat{G}(z'',z';p)\psi(z';p)\ , \tag{3}$$

or

$$\psi_j(\mathbf{x}_\perp'',z'';p)=\sum_k\int d^2\mathbf{x}_\perp'\langle\mathbf{x}_\perp''|\hat{G}_{jk}(z'',z';p)|\mathbf{x}_\perp'\rangle \times\psi_k(\mathbf{x}_\perp',z';p),\quad j,k=1,2,3,4\ , \tag{4}$$

for the time-Fourier component $\psi(\mathbf{x};p)$ of $\Psi(\mathbf{x},t)$. Then we have

$$\begin{aligned}\Psi(z'',t)&=\int_{p_0-\Delta p}^{p_0+\Delta p}dp\exp[-(i/\hbar)E(p)t]\hat{G}(z'',z';p)\psi(z';p)\\
&\approx\hat{G}(z'',z';p_0)\Psi(z',t)\ .\end{aligned}$$

(In the practically monoenergetic case, $\Delta p\approx 0$.) (5)

It is clear that the desired z propagator $\hat{G}(z'',z';p)$ for $\psi(z;p)$ defined through (3) is arrived at by integrating for z evolution the time-independent equation

$$\left\{E(p)+e\Phi-mc^2\beta-c\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp +c\alpha_z\left[i\hbar\frac{\partial}{\partial z}-\left[\frac{e}{c}\right]A_z\right]\right\}\psi=0\ , \tag{6}$$

obtained by substituting (2) in (1). Let us now rewrite this equation (6) in a more convenient form. We shall take

$$\Phi(\mathbf{x})=\Phi_0(z)+\phi(\mathbf{x}_\perp,z),\quad \Phi_0(z)=\Phi(0,0,z)\ , \tag{7}$$

such that $\Phi_0(z)$ represents the electric potential along the system axis and $\phi(\mathbf{x}_\perp,z)$ is the deviation in the potential in the off-axis region from $\Phi_0(z)$; further, without loss of generality, let us assume that $\Phi(\mathbf{x}_\perp,z<z_l)=\Phi(\mathbf{x}_\perp,z>z_r)=0$. Now, multiplying (6) by $(1/c)\alpha_z$ throughout from the left and rearranging the terms we get

$$\begin{aligned}
&i\hbar\frac{\partial\psi}{\partial z}=\hat{H}\psi\ ,\\
&\hat{H}=[(e/c)A_z-p(z)\beta\chi(z)\alpha_z-(e/c)\phi\alpha_z+\alpha_z\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp]\ ,\\
&p(z)=(1/c)\{[E(p)+e\Phi_0]^2-m^2c^4\}^{1/2}\ ,\\
&\chi(z)=\begin{bmatrix}\xi(z)I & 0\\ 0 & -\xi(z)^{-1}I\end{bmatrix},\\
&\xi(z)=[E(p)+e\Phi_0+mc^2]/cp(z)\ .
\end{aligned} \tag{8}$$

It may be noted that the expression $p(z)$ corresponds to the momentum of a classical electron moving in the one-dimensional region $z_l<z<z_r$ along the axis with the potential energy $-e\Phi_0(z)$ given that the initial momentum (at $z<z_l$) before entering the field region $(z>z_l)$ is p. Hereafter, expressions such as $\chi(z)$, $\xi(z)$, etc. will be simply denoted as χ, ξ, etc., respectively, without explicitly mentioning their z dependence, unless there is some necessity, or possibility of confusion.

It would be helpful to bring (8) to a form very close to (1), with z taking the place of t; this would facilitate adapting suitably the well-known methods of analyzing the time evolution of (1), to study the z evolution of (8). Hence, let us introduce the transformation

$$\psi\to\psi'=M\psi,\quad M=(1/\sqrt{2})(I+\chi\alpha_z)\ , \tag{9}$$

leading to the desired "beam-optical" representation of (8) [or (6)]:

$$i\hbar\frac{\partial\psi'}{\partial z}=\hat{H}'\psi' ,$$

$$\hat{H}'=\{M\hat{H}M^{-1}-i\hbar M[\partial(M^{-1})/\partial z]\}$$
$$=-p(z)\beta+(e/c)A_z-\{(e/2c)[\xi+(1/\xi)]\phi-(i\hbar/2\xi)d\xi/dz\}\beta$$
$$+\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp+\{(e/2c)[\xi-(1/\xi)]\phi+(i\hbar/2\xi)(d\xi/dz)\beta\}\chi\alpha_z . \quad (10)$$

With the motivation that will soon be clear we shall rewrite $\hat{H}'$ as

$$\hat{H}'=-p(z)\beta+\hat{\mathcal{E}}+\hat{\mathcal{O}} ,$$

$$\hat{\mathcal{E}}=(e/c)A_z-\{(e/2c)[\xi+(1/\xi)]\phi-(i\hbar/2\xi)d\xi/dz\}\beta$$

$$=\begin{bmatrix} ((e/c)A_z-\{[e/c^2p(z)][E(p)+e\Phi_0]\}\phi + (i\hbar/2\xi)d\xi/dz)I & 0 \\ 0 & ((e/c)A_z+\{[e/c^2p(z)][E(p)+e\Phi_0]\}\phi - (i\hbar/2\xi)d\xi/dz)I \end{bmatrix} ,$$

$$\hat{\mathcal{O}}=\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp+\{(e/2c)[\xi-(1/\xi)]\phi+(i\hbar/2\xi)(d\xi/dz)\beta\}\chi\alpha_z$$

$$=\begin{bmatrix} 0 & \xi\boldsymbol{\sigma}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp+\{[em\xi/p(z)]\phi + (i\hbar/2)d\xi/dz\}\sigma_z \\ -(1/\xi)\boldsymbol{\sigma}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp-\{[em/\xi p(z)]\phi - (i\hbar/2\xi^2)d\xi/dz\}\sigma_z & 0 \end{bmatrix} , \quad (11)$$

and note that $\beta\hat{\mathcal{E}}=\hat{\mathcal{E}}\beta$ and $\beta\hat{\mathcal{O}}=-\hat{\mathcal{O}}\beta$.

Some aspects of the standard Dirac equation, relevant for the problem on hand, may be recalled here. In the nonrelativistic situation any positive-energy Dirac spinor Ψ has the upper pair of components large compared to the lower pair. Any operator which couples the large to the small components, such as $\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}}$ in (1), is called "odd;" the operators which do not couple the large and small components are "even" [such as β, $e\Phi$, $\mathbf{S}=(\hbar/2)\begin{pmatrix}\sigma & 0\\ 0 & \sigma\end{pmatrix}$, etc.]. To analyze the Dirac equation as a nonrelativistic part and relativistic corrections the proper procedure is the Foldy-Wouthuysen method[11] which consists of reducing the strength of the odd operators in the equation systematically through a sequence of transformations such that in the transformed equation the Hamiltonian is effectively an even operator expressed as a series in terms of the expansion parameter $1/mc^2$ with the leading nonrelativistic part followed by relativistic correction terms (cf. also Ref. 12. For a general discussion of the role of the Foldy-Wouthuysen-type transformations in particle interpretations of relativistic wave equations cf. Ref. 13).

Coming back to (10) we find its similarity to (1) very striking. In the paraxial situation, with the transverse component of momentum very small compared to the axial component, any solution ψ' of (10) representing an electron beam propagating in the $+z$ direction would have the upper pair of components large compared to the lower pair; this can be easily verified, particularly in the field-free case. Hence, the separation of $\hat{H}'$ into odd and even operators as in (11), in analogy with the Dirac equation (1), should be helpful; the odd part $\hat{\mathcal{O}}$ will couple the large to the small components of ψ' and $[-p(z)\beta+\hat{\mathcal{E}}]$ is even. The next step to be followed is obvious now: we should reduce the strength of the odd operators in (10) systematically through a sequence of Foldy-Wouthuysen-like transformations such that in the new representation the right-hand side (rhs) of (10) is effectively even and given by a series expression in terms of the expansion parameter $1/p(z)$ with the leading paraxial part followed by nonparaxial (aberration) corrections calculated up to any desired order of accuracy.

To obtain the desired transformation of (10) we shall just adopt the Foldy-Wouthuysen prescription simply changing the algebraic expressions as suited to the present case. So, let

$$\psi'\rightarrow\psi_1=[\exp(i\hat{S}_1)]\psi', \quad \hat{S}_1=[i/2p(z)]\beta\hat{\mathcal{O}} . \quad (12)$$

We shall carry out the calculations up to the accuracy of the third power of the expansion parameter, $[1/p(z)]^3$, that would correspond to calculating up to the accuracy of the third-order aberrations in the classical treatment. Then, the transformation (12) takes (10) into the form

$$
\begin{aligned}
&i\hbar\frac{\partial\psi_1}{\partial z}=\hat{H}_1\psi_1 ,\\
&\hat{H}_1=\left[[\exp(i\hat{S}_1)]\hat{H}'[\exp(-i\hat{S}_1)]-i\hbar[\exp(i\hat{S}_1)]\left\{\frac{\partial[\exp(-i\hat{S}_1)]}{\partial z}\right\}\right]\\
&\quad=\hat{H}'-\hbar\left[\frac{\partial\hat{S}_1}{\partial z}\right]+i\left[\hat{S}_1,\hat{H}'-\tfrac{1}{2}\hbar\left[\frac{\partial\hat{S}_1}{\partial z}\right]\right]+(i^2/2!)\left[\hat{S}_1,\left[\hat{S}_1,\hat{H}'-\tfrac{1}{3}\hbar\left[\frac{\partial\hat{S}_1}{\partial z}\right]\right]\right]\\
&\quad\quad+(i^3/3!)\left[\hat{S}_1,\left[\hat{S}_1,\left[\hat{S}_1,\hat{H}'-\tfrac{1}{4}\hbar\left[\frac{\partial\hat{S}_1}{\partial z}\right]\right]\right]\right]+\cdots\\
&\quad=-p(z)\beta+\hat{\mathcal{E}}_1+\hat{\mathcal{O}}_1 ,\\
&\hat{\mathcal{E}}_1\approx\hat{\mathcal{E}}-[1/2p(z)]\beta\hat{\mathcal{O}}^2-[1/8p(z)^2]\left[\hat{\mathcal{O}},[\hat{\mathcal{O}},\hat{\mathcal{E}}]+i\hbar\left[\frac{\partial\hat{\mathcal{O}}}{\partial z}\right]\right]+[1/8p(z)^3]\beta\hat{\mathcal{O}}^4 ,\\
&\hat{\mathcal{O}}_1\approx-[1/2p(z)]\beta\left[[\hat{\mathcal{O}},\hat{\mathcal{E}}]+i\hbar\left[\frac{\partial\hat{\mathcal{O}}}{\partial z}\right]\right]+[i\hbar/2p(z)^2](dp/dz)\beta\hat{\mathcal{O}}-[1/3p(z)^2]\hat{\mathcal{O}}^3\\
&\quad\quad+[1/48p(z)^3]\beta\left[\hat{\mathcal{O}},\left[\hat{\mathcal{O}},[\hat{\mathcal{O}},\hat{\mathcal{E}}]+i\hbar\left[\frac{\partial\hat{\mathcal{O}}}{\partial z}\right]\right]\right] .
\end{aligned}
\tag{13}
$$

It is seen that while the term $\hat{\mathcal{O}}$ in $\hat{H}'$ is of order unity {i.e., $[1/p(z)]^0$}, the odd part $\hat{\mathcal{O}}_1$ in $\hat{H}_1$ contains only terms of order $1/p(z)$ and higher powers of $1/p(z)$; we shall write $\hat{\mathcal{O}}_1=O(1/p(z))$, and it is to be noted that ξ can be taken to be of order unity. Repeating the same type of transformation, now applied to (13), with

$$\psi_1\rightarrow\psi_2=[\exp(i\hat{S}_2)]\psi_1,\quad \hat{S}_2=[i/2p(z)]\beta\hat{\mathcal{O}}_1 , \tag{14}$$

leads to the result

$$
\begin{aligned}
&i\hbar\frac{\partial\psi_2}{\partial z}=\hat{H}_2\psi_2 ,\\
&\hat{H}_2=-p(z)\beta+\hat{\mathcal{E}}_2+\hat{\mathcal{O}}_2 ,
\end{aligned}
\tag{15}
$$

where the odd part $\hat{\mathcal{O}}_2=O([1/p(z)]^2)$; the explicit expressions for $\hat{\mathcal{E}}_2$ and $\hat{\mathcal{O}}_2$ can be obtained, respectively, from $\hat{\mathcal{E}}_1$ and $\hat{\mathcal{O}}_1$ through the replacements $\hat{\mathcal{E}}\rightarrow\hat{\mathcal{E}}_1$, $\hat{\mathcal{O}}\rightarrow\hat{\mathcal{O}}_1$ and calculations up to terms of order $[1/p(z)]^3$ as desired. After one more such transformation the resulting odd part is $O([1/p(z)]^3)$ which we shall neglect compared to the even parts retained up to $O([1/p(z)]^3)$.

The net result of the sequence of these transformations is as follows. Let

$$\psi'\rightarrow\psi_0=\hat{T}\psi' , \tag{16}$$

with

$$
\begin{aligned}
&\hat{T}=[\exp(i\hat{S}_3)][\exp(i\hat{S}_2)][\exp(i\hat{S}_1)] ,\\
&\hat{S}_1=[i/2p(z)]\beta\hat{\mathcal{O}} ,\\
&\hat{S}_j=[i/2p(z)]\beta\hat{\mathcal{O}}_{j-1},\quad j>1 ,\\
&\hat{\mathcal{E}}_j=\hat{\mathcal{E}}_1(\hat{\mathcal{E}}\rightarrow\hat{\mathcal{E}}_{j-1},\hat{\mathcal{O}}\rightarrow\hat{\mathcal{O}}_{j-1}),\quad j>1 ,\\
&\hat{\mathcal{O}}_j=\hat{\mathcal{O}}_1(\hat{\mathcal{E}}\rightarrow\hat{\mathcal{E}}_{j-1},\hat{\mathcal{O}}\rightarrow\hat{\mathcal{O}}_{j-1}),\quad j>1 ,
\end{aligned}
\tag{17}
$$

where $\hat{\mathcal{E}}_1(\hat{\mathcal{E}},\hat{\mathcal{O}})$ and $\hat{\mathcal{O}}_1(\hat{\mathcal{E}},\hat{\mathcal{O}})$ are given in (13) and the computations of $\{\hat{S}\}$ are carried out up to the accuracy of $O([1/p(z)]^3)$. Then, one has

$$
\begin{aligned}
&i\hbar\frac{\partial\psi_0}{\partial z}=\hat{H}_0\psi_0 ,\\
&\hat{H}_0\approx-p(z)\beta+\hat{\mathcal{E}}-[1/2p(z)]\beta\hat{\mathcal{O}}^2\\
&\quad-[1/8p(z)^2]\left[\hat{\mathcal{O}},[\hat{\mathcal{O}},\hat{\mathcal{E}}]+i\hbar\left[\frac{\partial\hat{\mathcal{O}}}{\partial z}\right]\right]\\
&\quad+[1/8p(z)^3]\beta\left\{\hat{\mathcal{O}}^4+\left[[\hat{\mathcal{O}},\hat{\mathcal{E}}]+i\hbar\left[\frac{\partial\hat{\mathcal{O}}}{\partial z}\right]\right]^2\right\} ,
\end{aligned}
\tag{18}
$$

up to the desired order of accuracy.

Let us first study (18) in the field-free input space where we choose Φ and $\mathbf{A}$ to be zero; this would help us fix the initial condition on ψ_0. With the subscript F indicating quantities belonging to the field-free space,

$$
\begin{aligned}
&p_F(z)\equiv p,\quad \xi_F=[E(p)+mc^2]/cp ,\\
&\chi_F=\begin{pmatrix}\xi_F I & 0\\ 0 & -\xi_F^{-1}I\end{pmatrix} ,\\
&\hat{\mathcal{E}}_F=0,\quad \hat{\mathcal{O}}_F=\chi_F\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp ,\\
&\hat{H}_{0F}\approx[-p+(1/2p)\hat{p}_\perp^2+(1/8p^3)\hat{p}_\perp^4]\beta\\
&\quad\approx-(p^2-\hat{p}_\perp^2)^{1/2}\beta=-\hat{\mathcal{P}}_z(p)\beta ,
\end{aligned}
\tag{19}
$$

and defining

$$\psi_0=\hat{\mathcal{T}}_0\psi,\quad \hat{\mathcal{T}}_0=\hat{T}\mathbf{M} , \tag{20}$$

we have

$$\hat{\mathcal{T}}_{0F}=\hat{T}_F M_F \approx (1/\sqrt{2})(\exp\{-(\beta\chi_F \boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp)[(1/2p)-(1/6p^3)\hat{p}_\perp^2]\})(I+\chi_F\alpha_z) . \quad (21)$$

Consequently, in the field-free input space

$$\psi_{0(\mathrm{in})}(z<z_l)=\hat{\mathcal{T}}_{0F}\psi_{(\mathrm{in})}(z<z_l) ,$$
$$i\hbar\frac{\partial\psi_{0(\mathrm{in})}}{\partial z}\approx -\hat{\mathcal{P}}_z\beta\psi_{0(\mathrm{in})} . \quad (22)$$

Since $\psi_{(\mathrm{in})}$ is a component of the forward-directed beam along the $+z$ axis it satisfies the relation

$$i\hbar\frac{\partial\psi_{(\mathrm{in})}}{\partial z}=-\hat{\mathcal{P}}_z\psi_{(\mathrm{in})} , \quad (23)$$

as can be easily verified from (2). The transformation operator $\hat{\mathcal{T}}_{0F}$ commutes with $i\hbar(\partial/\partial z)$ and $\hat{\mathcal{P}}_z$. Hence it follows that $\psi_{0(\mathrm{in})}$ also satisfies the same relation as (23):

$$i\hbar\frac{\partial\psi_{0(\mathrm{in})}}{\partial z}=-\hat{\mathcal{P}}_z\psi_{0(\mathrm{in})} . \quad (24)$$

For the consistency of (22) and (24) it must be true that

$$\beta\psi_{0(\mathrm{in})}\approx\psi_{0(\mathrm{in})} , \quad (25)$$

or in other words, up to the order of accuracy with which we are working, $\psi_{0(\mathrm{in})}$ must have the lower pair of components almost equal to zero (negligible) compared to the upper pair of components. Equation (25) specifies the required initial condition on ψ_0.

In general, we can write

$$\hat{H}_0=\begin{bmatrix}\hat{h}_1+\hat{h}_2 & 0\\ 0 & \hat{h}_1-\hat{h}_2\end{bmatrix} , \quad (26)$$

as can be seen by substituting the expressions for $\hat{\mathcal{E}}$ and $\hat{\mathcal{O}}$ given by (11), in (18). Here, $\hat{h}_1$ and $\hat{h}_2$ are 2×2 matrices and 0 is the 2×2 null matrix; for $\hat{H}_{0F}$, $\hat{h}_{1F}=0$ and $\hat{h}_{2F}=-\hat{\mathcal{P}}_z I$ as seen from (19). The z-evolution equation (18) for ψ_0, with the structure of $\hat{H}_0$ as given by (26), and the initial condition (25) imply that for all z, $\psi_0(z)$ has the lower pair of components negligible ($\approx$zero) compared to the upper pair, up to the desired order of accuracy, i.e., in general

$$\beta\psi_0(z)\approx\psi_0(z) . \quad (27)$$

In other words, in the ψ_0 representation one can employ, in practice, a two-component formalism to describe the beam propagation; this is exactly as in the case of the Foldy-Wouthuysen representation[11] of the Dirac electron theory. Instead, when the description of the system is desired to be fully in terms of the Dirac Ψ, as in (5), the significance of the relation (27) is that it greatly simplifies the computation of $\hat{G}$ as will be seen below.

In view of relation (27), we can write the ψ_0 equation (18) effectively as

$$i\hbar\frac{\partial\psi_0}{\partial z}\approx\hat{\mathcal{H}}_0\psi_0 ,$$
$$\hat{\mathcal{H}}_0=\begin{bmatrix}\hat{h}_1+\hat{h}_2 & 0\\ 0 & \hat{h}_1+\hat{h}_2\end{bmatrix} , \quad (28)$$

where $\hat{h}_1$ and $\hat{h}_2$ are as defined in (26) by $\hat{H}_0$. In the field-free region Eq. (28) becomes

$$i\hbar\frac{\partial\psi_{0F}}{\partial z}\approx\hat{\mathcal{H}}_{0F}\psi_{0F},\quad \hat{\mathcal{H}}_{0F}=-\hat{\mathcal{P}}_z , \quad (29)$$

as already noted in (24).

Now, the required expression for $\hat{G}(z'',z';p)$ follows immediately by integrating (28) formally and transforming back to the Dirac representation through relation (20). The result is as follows:

$$\hat{G}(z'',z';p)=[\hat{\mathcal{T}}_0(z'';p)]^{-1}\hat{G}_0(z'',z';p)[\hat{\mathcal{T}}_0(z';p)] ,$$
$$\hat{G}_0(z'',z';p)=P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}_0(z;p)\right]\right] , \quad (30)$$

where the symbol P stands for the z-ordered exponential. Consequently, for a practically monoenergetic beam of mean momentum p_0 the general propagation formula (5) takes the form

$$\Psi(z'',t)\approx\left\{[\hat{\mathcal{T}}_0(z'';p_0)]^{-1}P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}_0(z;p_0)\right]\right][\hat{\mathcal{T}}_0(z';p_0)]\right\}\Psi(z',t) . \quad (31)$$

Thus, we have arrived at a method of constructing the complete solution for the problem of determining the relation between the electron beam wave functions at different profile planes along the optic axis of the system, up to the desired order of accuracy.

The following are some general details useful in the application of the above formalism to particular situations.

(i) The explicit expression for the z-ordered exponential in (30) is given by

$$P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}_0(z)\right]\right]=\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}_0(z)+\tfrac{1}{2}(-i/\hbar)^2\int_{z'}^{z''}dz_1\int_{z'}^{z_1}dz[\hat{\mathcal{H}}_0(z_1),\hat{\mathcal{H}}_0(z)]+\cdots\right] , \quad (32)$$

following Magnus.[14] (For more details cf., e.g., Ref. 15.)

(ii) In view of the multiplicative property of the formal integration of any evolution equation like (28), the propagator $\hat{G}_0$ can be factorized as

$$\hat{G}_0(z'',z';p)=\hat{G}_0(z'',z;p)\hat{G}_0(z,z';p) , \tag{33}$$

if it is convenient to do so, with a suitable choice of z. Particularly, we shall use (33) to write

$$\hat{G}_0(z''>z_r,z'<z_l;p) =\hat{G}_{0F}(z'',z_r;p)\hat{G}_{0L}(z_r,z_l;p)\hat{G}_{0F}(z_l,z';p) , \tag{34}$$

where the subscripts F and L denote, respectively, free propagation and lens action. Consequently, we have

$$\hat{G}(z''>z_r,z'<z_l;p)=\hat{G}_F(z'',z_r;p)\hat{G}_L(z_r,z_l;p)\hat{G}_F(z_l,z';p) , \tag{35}$$

with

$$\hat{G}_F(z'',z';p)=\hat{\mathcal{T}}_{0F}^{-1}P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}_{0F}(p)\right]\right]\hat{\mathcal{T}}_{0F}$$
$$\equiv P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}_{0F}(p)\right]\right]=\hat{G}_{0F}(z'',z';p)=\exp[(i/\hbar)(z''-z')\hat{\mathcal{P}}_z(p)] \tag{36}$$

for all z',z'', and

$$\hat{G}_L(z_r,z_l;p)=\hat{\mathcal{T}}_{0F}^{-1}P\left[\exp\left[-(i/\hbar)\int_{z_l}^{z_r}dz\,\hat{\mathcal{H}}_0(z;p)\right]\right]\hat{\mathcal{T}}_{0F}=P\left[\exp\left[-(i/\hbar)\int_{z_l}^{z_r}dz\,\hat{\mathcal{H}}(z;p)\right]\right] , \tag{37}$$

defining

$$\hat{\mathcal{H}}(z;p)=\hat{\mathcal{T}}_{0F}^{-1}\hat{\mathcal{H}}_0(z;p)\hat{\mathcal{T}}_{0F} . \tag{38}$$

It may be noted that $\hat{\mathcal{H}}_F\equiv\hat{\mathcal{H}}_{0F}$.

(iii) We shall need the ψ_0 representation of a paraxial plane wave of momentum $(\mathbf{p}_\perp,p_z=+(p^2-p_\perp^2)^{1/2})$ with $|\mathbf{p}_\perp|\ll p$. In the Dirac representation the corresponding wave function has the form

$$\psi_{\mathbf{p}}=[\xi_F cp/2E(p)]^{1/2}\begin{pmatrix} a_+ \\ a_- \\ (a_+p_z+a_-p_-)/p\xi_F \\ (a_+p_+-a_-p_z)/p\xi_F \end{pmatrix}\exp\{(i/\hbar)[\mathbf{p}\cdot\mathbf{x}-E(p)t]\} ,$$
$$|a_+|^2+|a_-|^2=1 ,$$
$$p_\pm=p_x\pm ip_y,\quad |p_\pm|\ll p ,$$
$$p_z\approx p-(1/2p)p_\perp^2-(1/8p^3)p_\perp^4 , \tag{39}$$

with (a_+,a_-) being the amplitudes for spin-up and spin-down components along the $+z$ direction, respectively. The corresponding ψ_0 reads, as a result of (20) and (21),

$$\psi_{0\mathbf{p}}=\hat{\mathcal{T}}_{0F}\psi_{\mathbf{p}}$$
$$\approx(\{[1+(1/8p^2)\hat{p}_\perp^2+(11/128p^4)\hat{p}_\perp^4]$$
$$-[1+(3/8p^2)\hat{p}_\perp^2][(1/2p)\beta\chi_F\alpha_\perp\cdot\hat{\mathbf{p}}_\perp]\}[(1/\sqrt{2})(I+\chi_F\alpha_z)])\psi_{\mathbf{p}}$$
$$\approx[\xi_F cp/E(p)]^{1/2}\begin{pmatrix} \{[1-(3/8p^2)p_\perp^2-(13/128p^4)p_\perp^4]a_+ \\ +[(\tfrac{1}{2})-(1/16p^2)p_\perp^2](p_-/p)a_-\} \\ \{-[(\tfrac{1}{2})-(1/16p^2)p_\perp^2](p_+/p)a_+ \\ +[1-(3/8p^2)p_\perp^2-(13/128p^4)p_\perp^4]a_-\} \\ (31p_-/256\xi_F p^5)p_\perp^4 a_- \\ (31p_+/256\xi_F p^5)p_\perp^4 a_+ \end{pmatrix}\exp\{(i/\hbar)[\mathbf{p}\cdot\mathbf{x}-E(p)t]\}$$
$$\approx[\xi_F cp/E(p)]^{1/2}\begin{pmatrix} (1/2p)\{[2p_z(p+p_z)]^{1/2}a_+ \\ +[2p_z/(p+p_z)]^{1/2}p_-a_-\} \\ -(1/2p)\{[2p_z/(p+p_z)]^{1/2}p_+a_+ \\ -[2p_z(p+p_z)]^{1/2}a_-\} \\ 0 \\ 0 \end{pmatrix}\exp\{(i/\hbar)[\mathbf{p}\cdot\mathbf{x}-E(p)t]\} , \tag{40}$$

with the lower pair of components becoming zero (negligible) compared to the upper pair, up to the desired order of accuracy, as expected from (25).

(iv) In computing $\hat{\mathcal{H}}$ to obtain the lens propagator $\hat{G}_L$ the following results are useful. If we define

$$\hat{\mathbf{Q}}_\perp = \hat{\mathcal{T}}_{0F}^{-1}\mathbf{x}_\perp \hat{\mathcal{T}}_{0F} ,\quad \hat{\mathcal{S}} = \hat{\mathcal{T}}_{0F}^{-1}\mathbf{S}\hat{\mathcal{T}}_{0F} , \tag{41}$$

then

$$\begin{aligned}\hat{\mathbf{Q}}_\perp &\approx \mathbf{x}_\perp - i(\lambda\!\!\!^{-}/2)\beta\chi_F\boldsymbol{\alpha}_\perp[1+(1/2p^2)\hat{p}_\perp^2] \\ &\quad - i(\lambda\!\!\!^{-}/4p^2)\beta\chi_F(\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp)\hat{\mathbf{p}}_\perp + (\lambda\!\!\!^{-}/2\hbar p)(\mathbf{k}\times\hat{\mathbf{p}}_\perp)S_z ,\\ \hat{\mathcal{S}}_\perp &\approx -i(\lambda\!\!\!^{-}/2)[1+(1/2p^2)\hat{p}_\perp^2](\mathbf{k}\times\hat{\mathbf{p}}_\perp) \\ &\quad -[1+(1/4p^2)\hat{p}_\perp^2]\chi_F\alpha_z\mathbf{S}_\perp \\ &\quad +(1/2p^2)(\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp)\mathbf{S}_\perp(\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp) ,\\ \hat{\mathcal{S}}_z &\approx [1+(1/2p^2)\hat{p}_\perp^2][S_z-(1/p)\beta\chi_F\alpha_z\mathbf{S}_\perp\cdot\hat{\mathbf{p}}_\perp] ,\end{aligned} \tag{42}$$

with $\lambda\!\!\!^{-}=\hbar/p$ and $\mathbf{k}$ as the unit vector in the z direction, and as already noted,

$$\hat{\mathcal{T}}_{0F}^{-1}\hat{\mathbf{p}}_\perp\mathcal{T}_{0F}\equiv\hat{\mathbf{p}}_\perp . \tag{43}$$

In view of (42) and (43), $\hat{\mathcal{H}}$ defined by (38) can be obtained from $\hat{\mathcal{H}}_0$ by replacing $\mathbf{x}_\perp$ and $\mathbf{S}$, respectively, by $\hat{\mathbf{Q}}_\perp$ and $\hat{\mathcal{S}}$, and leaving $\hat{\mathbf{p}}_\perp$ unaltered:

$$\hat{\mathcal{H}}_0(\mathbf{x}_\perp,\hat{\mathbf{p}}_\perp,\mathbf{S},z;p)\rightarrow\hat{\mathcal{H}}=\hat{\mathcal{H}}_0(\mathbf{x}_\perp\rightarrow\hat{\mathbf{Q}}_\perp,\hat{\mathbf{p}}_\perp,\mathbf{S}\rightarrow\hat{\mathcal{S}},z;p) . \tag{44}$$

(v) It may be noted that the z evolution of the Dirac ψ governed by (8), or equivalently (31), is not unitary: $\int d^2x_\perp\psi^\dagger\psi$ is not necessarily constant along the axis. It is ψ_0 that has almost unitary z evolution: $\int d^2x_\perp\psi_0^\dagger\psi_0$ is almost conserved along the z axis.

III. EXAMPLES OF APPLICATION: PARAXIAL THEORY

We shall now consider some specific examples of application of the above general formalism of quantum theory of beam transport in an electron-lens system with a straight optic axis. In each case it is found that the well-known results based on the classical theory of electron optics are reproduced in the lowest-order approximation of the quantum theory.

The propagating electron beam will be considered to be not too intense so that the field due to the space charge can be neglected, and sufficiently paraxial ($|\pi_\perp| \ll p$) so that it is adequate to take only the first few terms in $\hat{\mathcal{H}}_0$ in (28), and accordingly simplify the expression for $\hat{\mathcal{T}}$ while computing $\hat{G}_0$ in (30). Further, in the ideal paraxial situation the effective field of the system influencing the beam dynamics is confined to a narrow vicinity of the axis so that the field potentials can be approximated by expressions containing only small powers of the off-axis coordinates $\mathbf{x}_\perp$. Thus, in the following illustrations we shall use the paraxial approximation which entails keeping in $\hat{\mathcal{H}}_0$ only terms up to quadratic in $(\mathbf{x}_\perp,\hat{\mathbf{p}}_\perp)$ and of order $(1/p)$ compared to the leading term, p. Also, we shall restrict the calculation to the first-order approximation of (32), keeping only the first term in the exponential. One has to go beyond these approximations to treat aberrating systems, and also to get quantum corrections to the classical results.

A. Axially symmetric magnetic lens

A magnetic lens with rotational symmetry about the z axis is formed by a static magnetic field $\mathbf{B}(\mathbf{x})$ for which the vector potential is usually chosen as

$$\mathbf{A}(\mathbf{x})=(-yA(\mathbf{x}),xA(\mathbf{x}),0) ,\quad A(x)=\tfrac{1}{2}B(z)-\frac{1}{16}\left[\frac{d^2B}{dz^2}\right]x_\perp^2+\frac{1}{384}\left[\frac{d^4B}{dz^4}\right]x_\perp^4-\cdots , \tag{45}$$

with the function $B(z)$ characterizing the field along the axis:

$$\mathbf{B}(0,0,z)=B(z)\mathbf{k},\quad z_l<z<z_r ,\quad B(z<z_l)=B(z>z_r)=0 . \tag{46}$$

In the paraxial situation $\mathbf{A}(\mathbf{x})$ is usually approximated as

$$\mathbf{A}(\mathbf{x})\approx(-\tfrac{1}{2}B(z)y,\tfrac{1}{2}B(z)x,0) . \tag{47}$$

There is no electric field associated with the system and so we shall choose $\Phi=0$. The other quantities of interest become

$$p(z)=p,\quad \xi(z)=\xi_F,\quad \chi(z)=\chi_F ,\quad \hat{\mathcal{E}}=0,\quad \hat{\mathcal{O}}=\chi_F\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp . \tag{48}$$

Then, using the paraxial approximation, we have

$$\begin{aligned}\hat{H}_0 &\approx -p\beta-(1/2p)\beta\hat{\mathcal{O}}^2 \\ &= \beta[-p+(1/2p)\hat{p}_\perp^2+(e^2/8c^2p)B(z)^2x_\perp^2 \\ &\quad +(e/2pc)B(z)(\hat{L}_z+2S_z)] ,\end{aligned} \tag{49}$$

where $\hat{L}_z$ is the z component of the angular momentum operator; correspondingly,

$$\begin{aligned}\hat{\mathcal{H}}_0 &\approx [-p+(1/2p)\hat{p}_\perp^2+(e^2/8c^2p)B(z)^2x_\perp^2 \\ &\quad +(e/2pc)B(z)(\hat{L}_z+2S_z)] .\end{aligned} \tag{50}$$

Since we are computing only up to the accuracy of $(1/p)$, under the paraxial approximation, the transformation operator $\hat{\mathcal{T}}_0$ can be taken to be

$$\hat{\mathcal{T}}_0\approx(1/\sqrt{2})[\exp(i\hat{S}_1)](I+\chi\alpha_z) ,\quad \hat{S}_1=(i/2p)\beta\hat{\mathcal{O}}=(i/2p)\beta\chi_F\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp ; \tag{51}$$

this is the operator which, in this case, transforms $\hat{H}$ in (8) into the $\hat{H}_0$ in (49) without any odd term up to the accuracy of $(1/p)$. Thus, in the case of an axially symmetric magnetic lens, we can see that for a practically monoenergetic incident beam with mean momentum p_0,

$$\Psi(z'',t)\approx\hat{G}(z'',z';p_0)\Psi(z',t) ,$$
$$\hat{G}(z'',z';p_0)=[\hat{\mathcal{T}}_0(z'';p_0)]^{-1}P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}_0(z;p_0)\right]\right][\hat{\mathcal{T}}_0(z';p_0)] , \tag{52}$$

where $\hat{\mathcal{H}}_0(z;p_0)$ and $\hat{\mathcal{T}}_0(z;p_0)$ are given, respectively, by (50) and (51) with p replaced by p_0; if V is the operating voltage of the system, through which the incident electrons are getting accelerated starting from rest, then, as is usually written,

$$p_0=(2emV^*)^{1/2}, \quad V^*=V(1+\epsilon V), \quad \epsilon=e/2mc^2 . \tag{53}$$

Now, we shall understand the focusing action of the system. To this end, let us compute the Dirac current density at the output plane at $z=z_r$ corresponding to the input wave function

$$\Psi_{(\mathrm{in})}(\mathbf{x}_\perp,z',t)=[\xi_F c p_0/2E(p)]^{1/2}\begin{pmatrix} a_+ \\ a_- \\ a_+/\xi_F \\ -a_-/\xi_F \end{pmatrix}\exp\{(i/\hbar)[p_0z'-E(p_0)t]\}, \quad z'<z_l ,$$
$$|a_+|^2+|a_-|^2=1 , \tag{54}$$

a plane wave of momentum p_0 in the z direction; the associated input current density is

$$\mathbf{j}_{(\mathrm{in})}(\mathbf{x}_\perp,z')=c[\Psi^\dagger_{(\mathrm{in})}(\mathbf{x}_\perp,z')]\boldsymbol{\alpha}[\Psi_{(\mathrm{in})}(\mathbf{x}_\perp,z')]=v(0,0,1), \quad z'<z_l , \tag{55}$$

where $v=c^2p_0/E(p_0)$ is the velocity of the beam electrons, and this corresponds clearly to a system of rays parallel to the z axis, the axis of the system.

In the case of the system under consideration, for any practically monoenergetic $\Psi_{(\mathrm{in})}(z'<z_l)$ corresponding to mean momentum p_0, the wave function in the output space becomes

$$\begin{aligned}\Psi_{(\mathrm{out})}(z'')&=\{[\hat{G}_F(z'',z_r;p_0)][\hat{G}_L(z_r,z_l;p_0][\hat{G}_F(z_l,z';p_0)]\}\Psi_{(\mathrm{in})}(z')\\ &\approx\{(\exp\{(i/\hbar)(z''-z_r)[p_0-(1/2p_0)\hat{p}_\perp^2]\})\\ &\quad\times[(\exp\{(i/\hbar)(z_r-z_l)[p_0-(1/2p_0)\hat{p}_\perp^2]-(i/2\lambda_0 f)x_\perp^2-(1/2f)\beta\chi_F(p_0)\mathbf{x}_\perp\cdot\boldsymbol{\alpha}_\perp\})\{\exp[-(i/\hbar)\theta(\hat{L}_z+S_z)]\}\\ &\quad\times(\exp\{-(i/\hbar)\theta[S_z-(1/p_0)\beta\chi_F(p_0)\alpha_z\mathbf{S}_\perp\cdot\hat{\mathbf{p}}_\perp]\})](\exp\{(i/\hbar)(z_l-z')[p_0-(1/2p_0)\hat{p}_\perp^2]\})\}\Psi_{(\mathrm{in})}(z') ,\end{aligned} \tag{56}$$

apart from a constant phase factor, where $\lambda_0=\hbar/p_0$,

$$\begin{aligned}1/f&=(e^2/4c^2p_0^2)\int_{z_l}^{z_r}dz\,B(z)^2\\ &\approx(e/8mc^2V^*)\int_{-\infty}^{\infty}dz\,B(z)^2 ,\end{aligned} \tag{57}$$

and

$$\theta=(e/2cp_0)\int_{z_l}^{z_r}dz\,B(z)=(e/8mc^2V^*)^{1/2}\int_{z_l}^{z_r}dz\,B(z) . \tag{58}$$

To arrive at relation (56) we have used (50)–(52) along with the results mentioned at the end of Sec. II [(32), (35)–(38), and (41)–(44)]; further, we have approximated the lens propagator $\hat{G}_L$ by its first-order expression taking only the first term in the exponential in (32).

The output wave function at the plane $z=z_r$ corresponding to the input wave function $\Psi_{(\mathrm{in})}$ in (54) is seen to be

$$\Psi_{(\mathrm{out})}(z_r)=\hat{G}_L(z_r,z_l)\Psi_{(\mathrm{in})}(z_l) , \tag{59}$$

as follows from (56). To compute the desired quantity $\mathbf{j}_{(\mathrm{out})}$ at $z''=z_r$ we shall assume that $z_r-z_l\ll f$. This enables us to separate the factors containing $\mathbf{x}_\perp$ and $\hat{\mathbf{p}}_\perp$ in $\hat{G}_L$ without regard to their noncommutativity. Then, up to first order in $\mathbf{x}_\perp$, it can be seen that

$$\begin{aligned}\mathbf{j}_{(\mathrm{out})}(\mathbf{x}_\perp,z_r)&=c[\Psi^\dagger_{(\mathrm{out})}(\mathbf{x}_\perp,z_r)]\boldsymbol{\alpha}[\Psi_{(\mathrm{out})}(\mathbf{x}_\perp,z_r)]\\ &\approx v(-x/f,-y/f,1) .\end{aligned} \tag{60}$$

Since the value of f, given by (57), is always positive Eq. (60) shows that the current density vector at every point in the transverse plane at $z=z_r$ is pointing towards the point $(0,0,z_r+f)$ on the axis, the focal point. In other words, the system acts as a converging lens of focal length f for the Dirac current. The assumption $z_r-z_l\ll f$ used here in computing $\mathbf{j}_{(\mathrm{out})}$ is precisely the usual "thin lens approximation." Thus, one obtains the celebrated classical formula of Busch[16] for the focal length of an axially symmetric thin magnetic electron lens as the lowest-order approximation result in the quantum theory based on the Dirac equation. Further, the presence of the term $\exp[-(i/\hbar)\theta\hat{L}_z]$ in (56) explains the quantum mechanics of the well-known image rotation through the angle θ given by (58).

The description of thin solenoidal magnetic lens under the paraxial approximation given here is the same as presented in our earlier paper;[1] there we have used the transformation technique of Sec. II with particular reference to the magnetic lens, whereas here the treatment has been derived as a special case of a general formalism applicable to any electron lens with straight axis. Further, in Ref. 1 we assume that $B(z)$ does not vary very much along the axis over a distance of the order of λ; this simplifies the transformations while going beyond the paraxial approximations, but does not make any difference at the level of paraxial theory.

B. Magnetic quadrupole lens

In this case the vector potential can be chosen to be given by

$$\mathbf{A}(\mathbf{x})=(0,0,\tfrac{1}{2}g_B(x^2-y^2)) , \tag{61}$$

as is usual, where g_B is a constant in the lens region (z_l,z_r) and zero outside. There is no electric field; we shall choose $\Phi=0$. The other quantities of interest are

$$\begin{aligned}
&p(z)=p,\quad \xi(z)=\xi_F,\quad \chi(z)=\chi_F ,\\
&\hat{\mathcal{E}}=(e/c)A_z,\quad \hat{\mathcal{O}}=\chi_F\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp\equiv\hat{\mathcal{O}}_F ,\\
&\hat{\mathcal{T}}_0\approx(1/\sqrt{2})\{\exp[-(1/2p)\beta\hat{\mathcal{O}}_F]\}(I+\chi_F\alpha_z)\approx\hat{\mathcal{T}}_{0F} .
\end{aligned} \tag{62}$$

Then, under the paraxial approximation,

$$\begin{aligned}
\hat{H}_0&\approx(e/c)A_z-\beta[p+(1/2p)\hat{\mathcal{O}}^2]\\
&=(e/c)A_z-\beta[p-(1/2p)\hat{p}_\perp^2] ,
\end{aligned} \tag{63}$$

and correspondingly

$$\begin{aligned}
\hat{\mathcal{H}}_0&\approx(e/c)A_z-p+(1/2p)\hat{p}_\perp^2\\
&=-p+(1/2p)\hat{p}_\perp^2+(e/2c)g_B(x^2-y^2) .
\end{aligned} \tag{64}$$

For a practically monoenergetic incident beam with mean momentum p_0 the relation between the input and the output wave functions becomes

$$\begin{aligned}
\Psi_{(\mathrm{out})}(z'')&=\{[\hat{G}_F(z'',z_r;p_0)][\hat{G}_L(z_r,z_l;p_0)][\hat{G}_F(z_l,z';p_0)]\}\Psi_{(\mathrm{in})}(z')\\
&\approx[\{\exp\{(i/\hbar)(z''-z_r)[p_0-(1/2p_0)\hat{p}_\perp^2]\}\}(\exp\{(i/\hbar)(z_r-z_l)[p_0-(1/2p_0)\hat{p}_\perp^2]-(i/2\lambda_0 f)(x^2-y^2)\\
&\qquad-(1/2f)\beta\chi_F(p_0)(x\alpha_x-y\alpha_y)\})\\
&\qquad\times(\exp\{(i/\hbar)(z_l-z')[p_0-(1/2p_0)\hat{p}_\perp^2]\})]\Psi_{(\mathrm{in})}(z') ,
\end{aligned} \tag{65}$$

where

$$1/f=k_B^2w,\quad k_B^2=eg_B/cp_0,\quad w=(z_r-z_l) . \tag{66}$$

Calculating the output current, as before, for an input plane wave of momentum p_0 in the z direction, it is found that the system, for $k_B^2>0$, has focusing action in the xz plane and defocusing action in the yz plane with the focal lengths

$$1/f_x=-1/f_y\approx1/f=k_B^2w , \tag{67}$$

up to first order in $\mathbf{x}_\perp$, as is well known in the classical treatment;[17] here, the thin lens condition is $w/f\ll1$, or $k_Bw\ll1$ as is usually written.

In this case, since $\hat{\mathcal{T}}_0\approx\hat{\mathcal{T}}_{0F}$ as mentioned in (62), generally the propagation for a practically monoenergetic beam follows the relation

$$\begin{aligned}
&\Psi(z'')\approx\hat{G}(z'',z';p_0)\Psi(z') ,\\
&\hat{G}(z'',z';p_0)=P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}(z;p_0)\right]\right] ,
\end{aligned} \tag{68}$$

with

$$\begin{aligned}
\hat{\mathcal{H}}(z;p_0)\approx&-p_0+(1/2p_0)\hat{p}_\perp^2+(e/2c)g_B(x^2-y^2)\\
&-i(e\lambda_0/2c)g_B\beta\chi_F(p_0)(x\alpha_x-y\alpha_y) ,
\end{aligned} \tag{69}$$

under the paraxial approximation.

C. Axially symmetric electrostatic lens

An axially symmetric electrostatic lens consists of an electrostatic field $\mathbf{E}(\mathbf{x})$ which can be derived, in general, from the potential

$$\begin{aligned}
\Phi(\mathbf{x}_\perp,z)&=\Phi_0(z)+\phi(\mathbf{x}_\perp,z)\\
&=\Phi_0(z)-\frac{1}{4}\left(\frac{d^2\Phi_0}{dz^2}\right)x_\perp^2+\frac{1}{64}\left(\frac{d^4\Phi_0}{dz^4}\right)x_\perp^4-\cdots ,
\end{aligned} \tag{70}$$

and there is no magnetic field; we shall choose $\mathbf{A}=0$ and let $\Phi_0(z)=0$ outside the lens region (z_l,z_r). With $p(z)$, $\xi(z)$, and $\chi(z)$ given by the same expressions as in (8),

$$\begin{aligned}
&\hat{\mathcal{E}}=-\beta\{[e/c^2p(z)][E(p)+e\Phi_0]\phi-(i\hbar/2\xi)d\xi/dz\} ,\\
&\hat{\mathcal{O}}=\chi\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp+\{[em/p(z)]\phi+(i\hbar/2\xi)(d\xi/dz)\beta\}\chi\alpha_z .
\end{aligned} \tag{71}$$

As before, we shall consider the incident beam to be a practically monoenergetic one with p_0 as the mean momentum of the constituent electrons. Let us assume that the change in the local potential $\Phi_0(z)$ along the axis over a distance of the order of λ_0 is very small compared to the accelerating voltage V [$p_0=(2emV^*)^{1/2}$ as in (53)]; V is constant in the system. This enables us to drop the $d\xi/dz$ terms of $\hat{\mathcal{E}}$ and $\hat{\mathcal{O}}$ in (71). Then,

$$\hat{H}_0 \approx -p(z)\beta + \hat{\mathscr{E}} - [1/2p(z)]\beta\hat{\mathscr{O}}^2$$

$$\approx \beta \left[-p(z) + [1/2p(z)]\hat{p}_\perp^2 + [e/4c^2p(z)][E(p)+e\Phi_0] \left[\frac{d^2\Phi_0}{dz^2} \right] x_\perp^2 \right] , \quad (72)$$

under the paraxial approximation which entails keeping only terms up to the quadratic in $(\mathbf{x}_\perp, \hat{\mathbf{p}}_\perp)$ and up to order $1/p$ compared to the leading term $-p\beta$. Correspondingly,

$$\hat{\mathscr{H}}_0 \approx -p(z) + [1/2p(z)]\hat{p}_\perp^2 + [e/4c^2p(z)][E(p)+e\Phi_0] \left[\frac{d^2\Phi_0}{dz^2} \right] x_\perp^2 . \quad (73)$$

In the present case

$$\hat{\mathscr{T}}_0(z;p) \approx (1/\sqrt{2})(\exp\{-[1/2p(z)]\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp\}) \times (I + \chi\alpha_z) , \quad (74)$$

and the general propagation formula for a practically monoenergetic beam of mean momentum p_0 would be given by (31) with $\hat{\mathscr{H}}_0$ as in (73), $\hat{\mathscr{T}}_0$ as in (74), and p_0 replacing p. Thus the relation between the input and the output wave functions now becomes

$$\Psi_{(\text{out})}(z'') = \{[\hat{G}_F(z'',z_r;p_0)][\hat{G}_L(z_r,z_l;p_0)][\hat{G}_F(z_l,z';p_0)]\}\Psi_{(\text{in})}(z')$$

$$\approx \left[(\exp\{(i/\hbar)(z''-z_r)[p_0 - (1/2p_0)\hat{p}_\perp^2]\}) \left[\exp\left\{ (i/\hbar)\left[\int_{z_l}^{z_r} dz\, p_0(z) - \frac{1}{2}\left[\int_{z_l}^{z_r} dz\, p_0(z)^{-1} \right] \hat{p}_\perp^2 \right] - (i/2\lambda\!\!\!^{-}_0 f)x_\perp^2 - (1/2f)\beta\chi_F(p_0)\mathbf{x}_\perp\cdot\boldsymbol{\alpha}_\perp \right\} \right] \times (\exp\{(i/\hbar)(z_l-z')[p_0-(1/2p_0)\hat{p}_\perp^2]\}) \right] \Psi_{(\text{in})}(z') \quad (75)$$

with

$$p_0(z) = (1/c)\{[E(p_0)+e\Phi_0]^2 - m^2c^4\}^{1/2} \quad (76)$$

and

$$1/f = (e/2c^2p_0) \times \int_{z_l}^{z_r} dz \left[[E(p_0)+e\Phi_0] \left[\frac{d^2\Phi_0}{dz^2} \right] \Big/ p_0(z) \right] . \quad (77)$$

Calculating the output current, as before, for an input plane wave of momentum p_0 in the z direction it can be seen that the system is a converging lens for Dirac electrons with the focal length f given by (77) up to first order and thin-lens approximations.

To identify (77) with the standard form of the classical result for the focal length of thin electrostatic electron lens under paraxial conditions[18] we shall note the following. Let us write

$$E(p_0) = mc^2 + eV , \quad (78)$$

assuming that the beam electrons have been accelerated through the operating voltage of the system V starting from rest. Including the constant voltage V the total potential along the axis can be written as

$$U(z) = V + \Phi_0(z) \quad (79)$$

such that

$$p_0(z) = [2emU(z)^*]^{1/2}, \quad U^* = U(1+\epsilon U) ; \quad (80)$$

it may be noted that $p_0(z)$ becomes constant p_0 outside the lens region. Now, using (78)–(80) in (77) we get

$$1/f = \tfrac{1}{4}(V^*)^{-1/2} \times \int_{z_l}^{z_r} dz \left[(1+2\epsilon U) \left[\frac{d^2U}{dz^2} \right] \Big/ (U^*)^{1/2} \right] , \quad (81)$$

the familiar expression for $1/f$ in the case of a thin relativistic electrostatic lens with the same constant potential V in the object (input) and the image (output) spaces.

As is well known, in the case of a lens with different constant potentials in the object and the image spaces V in the expression (81) for the image space focal length refers to its value in the image space. In the present context this can be understood as follows: the result (77) has been obtained by calculating the current at $z = z_r$ and hence p_0 occurring outside the integral [derived from $\lambda\!\!\!^{-}_0$ in the expression for $\hat{G}_L$ in (75)] refers to the output space with the consequence that V in (81) corresponds to its value in the output (image) space.

D. Electrostatic quadrupole lens

For the electrostatic quadrupole lens the scalar potential may be chosen as

$$\Phi(\mathbf{x}) = \tfrac{1}{2}g_E(x^2-y^2) , \quad (82)$$

with g_E as a constant in the lens region (z_l,z_r) and zero outside; so, in this case,

$$\Phi_0 = 0, \quad \phi = \tfrac{1}{2}g_E(x^2-y^2) , \quad (83)$$

and since there is no magnetic field we can choose $\mathbf{A}=0$.

The other quantities of interest become

$$P(z)=p,\quad \xi(z)=\xi_F,\quad \chi(z)=\chi_F\ ,$$
$$\hat{\mathcal{E}}=-[eE(p)/c^2p]\phi\beta\ ,$$
$$\hat{\mathcal{O}}=\chi_F\boldsymbol{\alpha}_\perp\cdot\hat{\mathbf{p}}_\perp+(em/p)\phi\chi_F\alpha_z\ , \tag{84}$$
$$\hat{\mathcal{T}}_0\approx\hat{\mathcal{T}}_{0F}\approx(1/\sqrt{2})\{\exp[-(1/2p)\beta\hat{\mathcal{O}}_F]\}(I+\chi_F\alpha_z)$$

Under the paraxial approximation, as in the case of (72), we get

$$\hat{H}_0\approx-p\beta+\hat{\mathcal{E}}-(1/2p)\beta\hat{\mathcal{O}}^2$$
$$\approx\beta\{-p+(1/2p)\hat{p}_\perp^2-[eE(p)/2c^2p]g_E(x^2-y^2)\}\ . \tag{85}$$

Correspondingly,

$$\hat{\mathcal{H}}_0\approx\{-p+(1/2p)\hat{p}_\perp^2-[eE(p)/2c^2p]g_E(x^2-y^2)\}\ . \tag{86}$$

For a practically monoenergetic beam of mean momentum p_0 the general propagation formula reads, under the paraxial approximation,

$$\Psi(z'')\approx\left\{P\left[\exp\left[-(i/\hbar)\int_{z'}^{z''}dz\,\hat{\mathcal{H}}(z;p_0)\right]\right]\right\}\Psi(z')\ ,$$
$$\hat{\mathcal{H}}(z;p_0)\approx-p_0+(1/2p_0)\hat{p}_\perp^2-[eE(p_0)/2c^2p_0]g_E(x^2-y^2)+[ie\lambda\!\!\!^{-}_0E(p_0)/2c^2p_0]g_E\beta\chi_F(p_0)(x\alpha_x-y\alpha_y)\ . \tag{87}$$

Now, the output wave function is given by

$$\Psi_{(\text{out})}(z'')=\{[\hat{G}_F(z'',z_r;p_0)][\hat{G}_L(z_r,z_l;p_0)][\hat{G}_F(z_l,z';p_0)]\}\Psi_{(\text{in})}(z')$$
$$\approx[(\exp\{(i/\hbar)(z''-z_r)[p_0-(1/2p_0)\hat{p}_\perp^2]\})$$
$$\times(\exp\{(i/\hbar)(z_r-z_l)[p_0-(1/2p_0)\hat{p}_\perp^2]-(i/2\lambda\!\!\!^{-}_0f)(x^2-y^2)-(1/2f)\beta\chi_F(p_0)(x\alpha_x-y\alpha_y)\})$$
$$\times(\exp\{(i/\hbar)(z_l-z')[p_0-(1/2p_0)\hat{p}_\perp^2]\})]\Psi_{(\text{in})}(z')\ . \tag{88}$$

with

$$1/f=k_E^2w,\quad k_E^2=-[eE(p_0)g_E/c^2p_0^2],\quad w=(z_r-z_l)\ . \tag{89}$$

Again, calculating as before the output current corresponding to an input plane wave of momentum p_0 in the z direction we rediscover the classical result: the system for $k_E^2>0$ has focusing action in the xz plane and defocusing action in the yz plane with the focal lengths

$$1/f_x=-1/f_y\approx1/f=k_E^2w\ , \tag{90}$$

up to first order in $\mathbf{x}_\perp$ and under the thin-lens approximation ($k_Ew\ll1$). To identify the expression for $1/f$ in (90) with the well-known result[17] it is enough to note that $E(p_0)/c^2p_0^2=1/p_0v$ where v is the velocity of the beam electrons.

IV. ON THE OPTICS OF DIRAC ELECTRONS IN SYSTEMS WITH CURVED OPTIC AXES

For studying the optics of electrons in a system with curved axis it is natural to start with the Dirac equation written in curvilinear coordinates adapted to the geometry of the system, or the classical design orbit. Following Ref. 19 let us choose the z axis (x^3 axis) coinciding with the space curve representing the optic axis of the system, or the ideal design orbit, and let the rectangular off-axis coordinates $\mathbf{x}^\perp=(x^1=x,\ x^2=y)$ at each z, be defined in such a way that the arc element ds is given by

$$ds^2=dx^2+dy^2+\zeta^2dz^2,\quad \zeta=(1-\mathbf{K}_\perp\cdot\mathbf{x}^\perp) \tag{91}$$

where $K_x(z)$ and $K_y(z)$ are the curvature components at z. [Reference 19 is concerned with micro electron-beam devices. In the context of accelerator optics, choice of coordinates for systems with curved optic axes is similar (for details cf. Ref. 20); in this case, there is no torsion in the axis and hence either $K_x=0$ and $K_y=0$.] Now, following Ref. 21 to construct the Dirac equation in a generally covariant form,[22] it is found that the required form of the Dirac equation can be written as

$$i\hbar\frac{\partial\Psi}{\partial t}=\left[mc^2\beta-e\Phi+c\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp+\zeta^{-1}c\alpha_z\left[-i\hbar\frac{\partial}{\partial z}+(e/c)\zeta A_z+\Gamma_z\right]\right]\Psi\ ,$$
$$\Gamma_z=(K_xS_y-K_yS_x)\ , \tag{92}$$

where

$$B_x=\zeta^{-1}\{[\partial(\zeta A_z)/\partial y]-(\partial A_y/\partial z)\}\ ,$$
$$B_y=\zeta^{-1}\{(\partial A_x/\partial z)-[\partial(\zeta A_z)/\partial x]\}\ ,$$

and

$$B_z=\{(\partial A_y/\partial x)-(\partial A_x/\partial y)\}\ .$$

To understand the evolution of the electron beam wave function along the curved optic axis of the system we can proceed from (92) exactly in the same way as in the case of the system with straight axis. As before, we can label the energies by the momenta $\mathbf{p}$ of the incident electrons and write

$$\Psi(\mathbf{x},t)=\int_{p_0-\Delta p}^{p_0+\Delta p} dp\, \exp[-(i/\hbar)E(p)t]\psi(\mathbf{x};p)\ ,$$
$$|\mathbf{p}_\perp| \ll p, \Delta p \ll p_0\ , \qquad (93)$$
$$E(p)=+(m^2c^4+c^2p^2)^{1/2}\ .$$

Then, analogous to (8) we have the time-independent equation for $\psi(\mathbf{x};p)$, the time-Fourier component of $\Psi(\mathbf{x},t)$, given by

$$i\hbar\frac{\partial\psi}{\partial z}(\mathbf{x}^\perp,z)=\hat{H}\psi(\mathbf{x}^\perp,z)\ ,$$
$$\hat{H}=[(e/c)\zeta A_z-\zeta p(z)\beta\chi\alpha_z-(e/c)\zeta\phi\alpha_z + \zeta\alpha_z\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp+\Gamma_z]\ , \qquad (94)$$

with the definitions of ϕ, $p(z)$, and χ being the same as given in (7) and (8). Now, we should get an expansion of the rhs of (94) as a power series in $\hat{\pi}_\perp/p(z)$ so that the system can be studied under suitable approximations. To this end, first we shall make the transformation

$$\psi\rightarrow\psi'=M\psi\ , \qquad (95)$$

with the same M as in (9). This transformation takes (94) into the desired beam-optical representation

$$i\hbar\frac{\partial\psi'}{\partial z}=\hat{H}'\psi'\ ,$$
$$\hat{H}'=-p(z)\zeta\beta+\hat{\mathcal{E}}+\hat{\mathcal{O}}\ ,$$
$$\hat{\mathcal{E}}=(e/c)\zeta A_z-\{[e/c^2p(z)][E(p)+e\Phi_0]\zeta\phi -(i\hbar/2\xi)d\xi/dz\}\beta\ , \qquad (96)$$
$$\hat{\mathcal{O}}=\zeta\chi\boldsymbol{\alpha}_\perp\cdot[\hat{\boldsymbol{\pi}}_\perp-(i\hbar/2\zeta)\mathbf{K}_\perp] +\{[em/p(z)]\zeta\phi+(i\hbar/2\xi)(d\xi/dz)\beta\}\chi\alpha_z\ ,$$

where $\hat{\mathcal{O}}$ is the odd part of $\hat{H}'$. Then, through a sequence of Foldy-Wouthuysen-like transformations,

$$\psi'\rightarrow\psi_1=[\exp(i\hat{S}_1)]\psi',\quad \hat{S}_1=[i/2\zeta p(z)]\beta\hat{\mathcal{O}}\ ,$$
$$i\hbar\frac{\partial\psi_1}{\partial z}=\hat{H}_1\psi_1,\quad \hat{H}_1=-p(z)\zeta\beta+\hat{\mathcal{E}}_1+\hat{\mathcal{O}}_1\ ,$$
$$\psi_1\rightarrow\psi_2=[\exp(i\hat{S}_2)]\psi_1,\quad \hat{S}_2=[i/2\zeta p(z)]\beta\hat{\mathcal{O}}_1\ , \qquad (97)$$
$$i\hbar\frac{\partial\psi_2}{\partial z}=\hat{H}_2\psi_2,\quad \hat{H}_2=-p(z)\zeta\beta+\hat{\mathcal{E}}_2+\hat{\mathcal{O}}_2,\ldots,$$
$$\psi_j\rightarrow\psi_{j+1}=[\exp(i\hat{S}_{j+1})]\psi_j,\quad \hat{S}_{j+1}=[i/2\zeta p(z)]\beta\hat{\mathcal{O}}_j\ ,$$
$$i\hbar\frac{\partial\psi_{j+1}}{\partial z}=\hat{H}_{j+1}\psi_{j+1},\quad \hat{H}_{j+1}=-p(z)\zeta\beta+\hat{\mathcal{E}}_{j+1} +\hat{\mathcal{O}}_{j+1},\ldots,$$

one can get $\{H_j\}$ with successively smaller odd parts ($\hat{\mathcal{O}}_{j+1}$ is smaller than $\hat{\mathcal{O}}_j$); in the course of these transformations ζ can be treated as a constant, i.e., $[(\hbar/p(z))\boldsymbol{\nabla}_\perp,\zeta]\approx 0$, for all practical purposes, since $|\hbar\mathbf{K}_\perp|$ is too small compared to $p(z)$ or in other words the de Broglie wavelength of the beam electrons is too small compared to the radii of curvatures involved. Stopping at any desired step of the ladder of transformations (97) we can write, as before,

$$i\hbar\frac{\partial\psi_0}{\partial z}\approx\hat{H}_0\psi_0,\quad \psi_0=\hat{T}M\psi\ , \qquad (98)$$

where $\hat{H}_0$ is completely even up to the required level of accuracy. Here again, we would have $\beta\psi_0\approx\psi_0$, since the beam should be always forward-directed along the optic axis; consequently, we can replace (98) by

$$i\hbar\frac{\partial\psi_0}{\partial z}\approx\hat{\mathscr{H}}_0\psi_0\ , \qquad (99)$$

where $\hat{\mathscr{H}}_0$ would have only identical 2×2 blocks along the diagonal. Following (99) specific systems can be studied using the same techniques employed in treating straight-axis systems in previous sections.

We shall close this section with the following observation: For a purely magnetic system it is easy to see that

$$\hat{\mathscr{H}}_0\approx -p(z)\zeta+(e/c)\zeta A_z+[1/2p(z)]\zeta\hat{\pi}_\perp^2+\cdots\ , \qquad (100)$$

which goes over to the geometrical optics Hamiltonian in the classical limit (cf. Ref. 20 for details of the classical treatment).

V. CONCLUDING REMARKS

To summarize, following our previous paper[1] on solenoidal magnetic lenses, here I have developed the quantum theory of general electron lens systems entirely on the basis of the Dirac equation; straight optic axis systems have been dealt with in detail and extension of the theory to systems with curved axes has been discussed briefly. I hope to return to the detailed study of specific curved axis systems elsewhere.

As examples of application of the general theory in the case of straight axis systems the solenoidal magnetic lens, the quadrupole magnetic lens, the axially symmetric electrostatic lens, and the electrostatic quadrupole lens have been studied. In all these cases, the results of the classical (geometrical optics) theory in the paraxial approximation have been rediscovered naturally in the lowest-order approximation of the quantum theory.

The way in which the quantum theory approaches the classical theory is as follows. The $\hat{\mathscr{H}}_0$ appearing in the ψ_0 equation becomes the geometrical optics Hamiltonian (relativistic, including aberrations) upon "classicalization" which entails dropping the nonclassical (or, matrix) terms, replacing the quantum operators $(\hat{\mathbf{x}}^\perp,\hat{\mathbf{p}}_\perp)$ by the classical canonical variables $(\mathbf{x}^\perp,\mathbf{p}_\perp)$, and letting $\hbar\rightarrow 0$ wherever necessary; the nonrelativistic theory follows from the approximation $E(p)-mc^2\ll mc^2$. Correspondingly, one can see that the Heisenberg-type equations of motion for $(\mathbf{x}^\perp(z),\hat{\mathbf{p}}_\perp(z))$ (z taking the role of t) with the scalar part of $\hat{\mathscr{H}}_0$ as the Hamiltonian reproduce the Hamiltonian equations, of geometrical optics, for $(\mathbf{x}^\perp(z),\mathbf{p}_\perp(z))$. Thus, it is clear that, to obtain the quantum corrections to the classical theory, one should not ignore the essential four-component nature of the Dirac electron wave function, and should go beyond the paraxial and the first-order approximations; quantum effects are really in the form of nonclassical aberrations. For example, the following is seen readily: the expression for $\hat{\mathbf{Q}}_{\perp}$ in (42) shows that when one takes into account terms of order $1/p^2$ and higher in forming $\hat{\mathcal{H}}$ in (44) the coefficients of $\hat{p}_{\perp}^2$ (free propagation term) and $\hat{p}_{\perp}^4$ (spherical aberration term) would get modified from their respective classical values; the spin-dependent terms are derived automatically and terms nonexistent in the geometrical optics Hamiltonian, like the analogue of the Darwin term of the Dirac theory, appear naturally in this treatment. Hence, it should be worthwhile to develop the theory of electron-optical aberrations based on the quantum theory derived from the Dirac equation as detailed above and explore how far the quantum effects are of significance from the point of view of practical design of electron-beam devices.

Let us now look at the scalar approximation of the above formalism at the level of quantum theory itself. Let us use the notation (Ψ) for a four-component wave function to distinguish it from a scalar Ψ. The Dirac (Ψ) representation and the (Ψ_0) representation defined in Sec. II are completely equivalent, related by an invertible transformation. We worked in the (Ψ) representation in Sec. III, but one can as well work entirely within the framework of the (Ψ_0) representation. Since (Ψ_0) has the lower pair of components small compared to the upper pair, the resulting theory is effectively a two-component formalism as already noted in Sec. II. To see what this leads to, let us find the expression for the current density in the (Ψ_0) representation: in the field-free region it is found that we can write

$$\begin{aligned}\mathbf{j}&=c(\Psi)^{\dagger}\boldsymbol{\alpha}(\Psi)/\int d^3x\,(\Psi)^{\dagger}(\Psi)\\&=c(\hat{\mathcal{T}}_0^{-1}\Psi_0)^{\dagger}\boldsymbol{\alpha}(\hat{\mathcal{T}}_0^{-1}\Psi_0)/\int d^3x(\Psi_0)^{\dagger}(\hat{\mathcal{T}}_0^{-1})^{\dagger}(\hat{\mathcal{T}}_0^{-1})(\Psi_0)\\&\approx\mathrm{Re}\left[c^2(\Psi_0)^{\dagger}\hat{\mathbf{p}}(\Psi_0)/E(p)\int d^3x(\Psi_0)^{\dagger}(\Psi_0)\right],\end{aligned}\qquad(101)$$

up to first order in $\hat{\mathbf{p}}$. This implies that if we had worked in the (Ψ_0) representation in Sec. III, without transforming back to the (Ψ) representation following (30), we would have obtained the same results for the focusing action of the lenses, such as (60), using the expression

$$\mathbf{j}\approx\mathrm{Re}[c^2(\Psi_0)^{\dagger}\hat{\mathbf{p}}(\Psi_0)/E(p)]\,,\qquad(102)$$

for the current density associated with a normalized (Ψ_0). Now, let us note that $E(p)/c^2$ in (102) is just the relativistic mass, and in the nonrelativistic limit $E(p)/c^2\approx m$. Hence, in the nonrelativistic situation $\{p\approx[2m(E-mc^2)]^{1/2}\}$ if we approximate the formalism based on the (Ψ_0) representation to a scalar theory, by dropping all the nonscalar matrix terms in $\hat{\mathcal{H}}_0$, then the corresponding approximation of (102) reads

$$\mathbf{j}\approx\mathrm{Re}(\Psi_0^*\hat{\mathbf{p}}\Psi_0/m)\,,\qquad(103)$$

identical to the formula for the current density associated with a nonrelativistic Schrödinger wave function Ψ_0. Thus, it is clear that the scalar approximation of (Ψ_0), or Ψ_0, can be identified with the Schrödinger wave function in the nonrelativistic case. This nonrelativistic Ψ_0 may be called the Glaser wave function in the context of electron optics. Actually, in the nonrelativistic paraxial scalar approximation, the ψ_0-equation (28) is identical to Glaser's "Schrödinger equation for paraxial electron motion" (Eq. 45.10 of Ref. 2). Glaser obtains this equation by quantizing directly the classical nonrelativistic paraxial trajectory equation (cf. Ref. 23 for analogous wavization of geometrical light optics). Another way to obtain the Glaser equation is to subject the nonrelativistic Schrödinger equation directly to a paraxial approximation, along the lines of the Leontovich-Fock treatment[24,25] of the Helmholtz equation in the case of light optics.

When the scalar approximation of the (Ψ_0) representation is carried over to the relativistic domain of p [i.e., $p=(E^2-m^2c^4)^{1/2}/c$] the resulting ψ_0 equation (28) is, in the paraxial case, identical to the one that is obtained by subjecting the Klein-Gordon equation to the Leontovich-Fock-type paraxial approximation. In such a relativistic scalar-Ψ_0 approximation logically one has to use just the scalar approximation of (102), namely,

$$\mathbf{j}\approx\mathrm{Re}[c^2\Psi_0^*\hat{\mathbf{p}}\Psi_0/E(p)]\,,\qquad(104)$$

for computing the current density associated with Ψ_0. Obviously, Eq. (104) can be viewed as the relativistic extension of the nonrelativistic Schrödinger formula (103), obtained by substituting the relativistic mass for the rest mass. In the nonrelativistic scalar approximation equations corresponding to (56), (65), (75), etc., for example, contain the de Broglie wavelength in the form h/mv. In the relativistic scalar approximation the de Broglie wavelength in these equations has the form h/p, with p as the relativistic momentum, as if the nonrelativistic λ has been replaced by the relativistically corrected expression $\lambda(1-v^2/c^2)^{1/2}$. From these considerations we can understand, in a way, the apparent success of the customary substitution rules used to extend the nonrelativistic quantum theory of electron optics into the relativistic domain, and the replacement of such rules by the unambiguous procedure of the FHS formalism based on the Klein-Gordon equation seems only natural. Thus, the relativistic scalar Ψ_0, obeying (28) approximately, with $\hat{\mathcal{H}}_0$ containing only the scalar terms, is to be identified with the Klein-Gordon wave function, and in the context of electron optics it may be called the FHS wave function. The configuration space matrix element $\langle\mathbf{x}_{\perp}|\hat{G}(z,z')|\mathbf{x}'_{\perp}\rangle$ of the scalar $\hat{G}$ is the kernel of the integral relation connecting the wave functions in the planes at z and z' along the optic axis, and in the Glaser and the FHS formalisms semiclassical methods lead to an evaluation of this kernel in terms of classical paths.

Now that we have the theoretical framework based entirely on the Dirac equation one can analyze the finer details of the behavior of electron-lens systems with better accuracy than it has been possible so far. To see that there will really be a difference in the understanding of the electron optical phenomena when the proper four-component Dirac theory is used instead of a scalar approximation thereof, it is enough to observe the following. In the scalar theory the focusing action of a thin electron lens in the paraxial situation is represented essentially by the relation

$$\Psi_{(\text{out})} \sim \{\exp[-i(x^2 \pm y^2)/2\bar{\lambda} f]\}\Psi_{(\text{in})} , \tag{105}$$

where $\pm$ correspond respectively to axially symmetric and quadrupole lenses; this is clear from the discussion in Sec. III when scalar approximation is considered. One may recall that in the scalar wave theory of light optics also it is the multiplication by the phase factor $\exp(-i\pi x_\perp^2/\lambda f)$ that represents essentially the action of a thin lens on the wave function of a paraxial monochromatic beam (cf., e.g., Refs. 26 and 27). The explanation of the electron-lens action is, then, the result of application of the current density formula (103) (in the nonrelativistic case), or (104) (in the relativistic case), to (105). In the Dirac theory Eq. (105) gets replaced by

$$(\Psi)_{(\text{out})} \sim \{\exp[-\beta\chi_F(x\alpha_x \pm y\alpha_y)/2f]\} \times \{\exp[-i(x^2 \pm y^2)/2\bar{\lambda} f]\}(\Psi)_{(\text{in})} , \tag{106}$$

as obtained in Sec. III. Now, it is found that the application of the Dirac current density formula $\mathbf{j} = c(\Psi)^\dagger \boldsymbol{\alpha}(\Psi)$ to (106) explains the lens action with the result depending solely on the presence of the matrix factor in (106), and completely independent of the scalar phase factor. Also, curiously, the matrix factor in (106) does not contain $\hbar$ at all! It should be noted that for any observable, represented by a quantum operator, say, $\hat{A}(\hat{\mathbf{x}}_\perp, \hat{\mathbf{p}}_\perp)$ the Dirac theory, with the extra matrix factors as in (106), would lead to an average value different from the corresponding value obtained in the approximate scalar theory. Hence, in the development of the quantum theory of aberrations in electron optics it should make a difference whether we use the Dirac spinor theory or a scalar approximation thereof (for complete details of the classical theory of aberrations in electron optics, cf. Ref. 28).

In the ultrarelativistic situation (very-high-energy accelerator optics), $v \approx c$, and Eq. (106) becomes

$$(\Psi)_{(\text{out})} \sim \{\exp[-(x\alpha_x - y\alpha_y)/2f]\} \times \{\exp[-i(x^2 - y^2)/2\bar{\lambda} f]\}(\Psi)_{(\text{in})} , \tag{107}$$

in the case of quadrupole lenses. Now, the focusing action of these lenses follows from the simple algebra

$$\{\exp[-(x\alpha_x - y\alpha_y)/2f]\}^\dagger \boldsymbol{\alpha}\{\exp[-(x\alpha_x - y\alpha_y)/2f]\} \approx (\alpha_x - (x/f), \alpha_y + (y/f), \alpha_z) , \tag{108}$$

up to first order in (x,y), independent of the scalar phase factor $\exp[-i(x^2 - y^2)/2\bar{\lambda} f]$.

There has always been a close analogy between electron optics and light optics. Our previous paper[1] and the present work are the result of looking for a convenient formulation of electron optics where the spin and the spatial degrees of freedom of the electron are treated in a uniform way consistent with the subtle manner in which the Dirac equation couples the spinor components, very much like the recent development[29] of a systematic procedure for handling the spatial variation and polarization of light in a uniform way consistent with the Maxwell equations. In fact, the formula (44) for getting the propagator $\hat{G}(z'' > z_r, z' < z_l)$ from the corresponding $\hat{G}_0$ is, except for the spin terms, exactly like the Mukunda-Simon-Sudarshan rule[29] for the passage from the scalar (Helmholtz) to the vector (Maxwell) wave theory in light optics; also, the above-mentioned difference between the scalar and the spinor treatments of electron optics is exactly similar to the difference between the scalar and the vector treatments of light wave optics (cf. Ref. 29 for details).

Recently, Lie algebraic techniques have been developed to clarify several issues in light optics, related to scalar wave optics,[27] Maxwell optics,[29] and ray optics.[30] These Lie algebraic techniques have also been adopted to analyze the problems of classical theory of charged particle optics, related to electron optical imaging process and accelerator systems[31] (cf. Ref. 32 for the early use of algebraic techniques in connection with accelerator optics). Now, the expression for the propagator $\hat{G}$ in the form of an exponential operator as in (30) shows that such Lie algebraic techniques retain their relevance in the quantum theory of electron optics as well. Particularly, the recent developments[31] in the application of Lie algebraic techniques to the classical theory of charged particle optics could be of much use, when suitably adapted to the quantum theory, to study aberrating systems following the operator approach presented here.

ACKNOWLEDGMENTS

I am very much grateful to Professor E. C. G. Sudarshan for proposing the problem, discussions, and a critical reading of the original draft of the manuscript. I wish to thank also Professor N. Mukunda and Dr. R. Simon, for helpful discussions on the topic of the paper. I would like to thank Professor N. R. Ranganathan for helping me understand Ref. 7. I am also indebted to Professor H. A. Ferwerda for certain useful remarks, in private correspondence, essentially leading to the discussion surrounding Eqs. (102)–(108). I am thankful to Professor P. W. Hawkes for making available to me a copy of the thesis mentioned in Ref. 7.

[1]R. Jagannathan, R. Simon, E. C. G. Sudarshan, and N. Mukunda, Phys. Lett. A **134**, 457 (1989).

[2]W. Glaser, in Vol. 33 of *Handbuch der Physik*, edited by S. Flugge (Springer-Verlag, Berlin, 1956).

[3]P. W. Hawkes, in *Image Processing and Computer-aided Design in Electron Optics*, edited by P. W. Hawkes (Academic, New York, 1973); in *Computer Processing of Electron Microscope Images*, Vol. 13 of *Topics in Current Physics*, edited by P. W. Hawkes (Springer-Verlag, Berlin, 1980).

[4]Cf., e.g., R. D. Heidenreich, *Fundamentals of Transmission Electron Microscopy* (Interscience, New York, 1964); F. A. Lenz, in *Electron Microscopy in Materials Science*, edited by U. Valdre (Academic, New York, 1971).

[5]Cf. K. Fujiwara, J. Phys. Soc. Jpn. **16**, 2226 (1961).

[6]A. Rubinowicz, Acta Phys. Polonica **3**, 143 (1934); in *Progress in Optics*, edited by E. Wolf (North Holland, Amsterdam, 1965), Vol. IV.

[7]Phan Van Loc, C. R. Acad. Sci. Paris **237**, 649 (1953); **238**, 2494 (1954); **246**, 388 (1958); D. Sc. thesis, University of Toulouse, 1960.

[8]Cf., e.g., T. Mulvey, *Magnetic Electron Lenses*, Vol. 18 of *Topics in Current Physics* (Springer-Verlag, Berlin, 1982); B. K. Vainshtein, Sov. Phys.—Usp. **30**, 393 (1987).

[9]H. A. Ferwerda, B. J. Hoenders, and C. H. Slump, Opt. Acta **33**, 145 (1986); **33**, 159 (1986).

[10]Cf., e.g., *Frontiers of Particle Beams, Observation, Diagnosis, and Correction*, Vol. 343 of *Lecture Notes in Physics*, edited by M. Month and S. Turner (Springer-Verlag, Berlin, 1989), and references therein.

[11]L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).

[12]S. Tani, Prog. Theor. Phys. **6**, 267 (1951).

[13]R. Acharya and E. C. G. Sudarshan, J. Math. Phys. **1**, 532 (1960).

[14]W. Magnus, Comm. Pure Appl. Math. **7**, 649 (1954).

[15]R. M. Wilcox, J. Math. Phys. **4**, 962 (1967); G. Dattoli, J. C. Gallardo, and A. Torre, Riv. Nuovo Cimento **11**, No. 11 (1988).

[16]H. Busch, Ann. Phys. (Leipzig) **81**, 974 (1926).

[17]Cf., e.g., H. Wollnik, *Optics of Charged Particles* (Academic, New York, 1987), Chap. 3.

[18]Cf., e.g., P. Grivet, *Electron Optics: Part I—Optics*, 2nd ed. (Pergamon, Oxford, 1972), Chap. 4; A. Zhigarev, *Electron Optics and Electron-Beam Devices* (Mir, Moscow, 1975), Chap. 1.

[19]Ximen Jiye, Optik **80**, 55 (1988); P. W. Hawkes and E. Kasper, *Principles of Electron Optics* (Academic, London, 1989), Vol. 2, Chap. 51.

[20]E. D. Courant and H. S. Snyder, Ann. Phys. **3**, 1 (1953); and, e.g., H. C. Corben and P. Stehle, *Classical Mechanics* (Wiley, New York, 1960), Chap. 17; also, *Optics of Charged Particles* (Ref. 17), Chap. 2.

[21]D. R. Brill and J. A. Wheeler, Rev. Mod. Phys. **29**, 465 (1957).

[22]Also, cf., E. Cartan, *The Theory of Spinors* (Hermann, Paris, 1966) Chap. IX; S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Oxford University Press, Oxford, 1983), Chap. 10; and, e.g., E. A. Lord, *Tensors, Relativity and Cosmology* (Tata McGraw Hill, New Delhi, 1976), Chap. 13.

[23]D. Marcuse, *Light Transmission Optics* (Van Nostrand, New York, 1972), Chap. 3; V. I. Man'ko and K. B. Wolf, in *Lie Methods in Optics*, Vol. 250 of *Lecture Notes in Physics*, edited by J. S. Mondragon and K. B. Wolf (Springer-Verlag, Berlin, 1986); T. Pradhan, Phys. Lett. A **122**, 397 (1987).

[24]M. Leontovich and V. Fock, Sov. J. Phys. **10**, 13 (1946).

[25]Also, cf., J. A. Arnaud, *Beam and Fiber Optics* (Academic, New York, 1976), Chap. 3; V. I. Man'ko, in *Lie Methods in Optics*, Vol. 250 of *Lecture Notes in Physics*, edited by J. S. Mondragon and K. B. Wolf (Springer-Verlag, Berlin, 1986).

[26]J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968), Chap. 5.

[27]H. Bacry and M. Cadilhac, Phys. Rev. **A26**, 2533 (1981); E. C. G. Sudarshan, R. Simon, and N. Mukunda, Phys. Rev. **A28**, 2921 (1983); H. Bacry, in *XIII International Colloquium on Group Theoretical Methods in Physics*, edited by W. W. Zachary (World Scientific, Singapore, 1984); A. J. Dragt, E. Forest, and K. B. Wolf, in *Lie Methods in Optics*, Vol. 250 of *Lecture Notes in Physics*, edited by J. S. Mondragon and K. B. Wolf (Springer-Verlag, Berlin, 1986); K. B. Wolf, Ann. Phys. **172**, 1 (1986).

[28]P. W. Hawkes, Optik **83**, 104 (1989); **83**, 122 (1989); P. W. Hawkes and E. Kasper, *Principles of Electron Optics* (Academic, London, 1989), Vol. 1, Part IV.

[29]N. Mukunda, R. Simon, and E. C. G. Sudarshan, Phys. Rev. A **28**, 2933 (1983); J. Opt. Soc. Am. A **2**, 416 (1985); R. Simon, E. C. G. Sudarshan, and N. Mukunda, *ibid.* **3**, 356 (1986); Appl. Opt. **26**, 1589 (1987).

[30]M. Nazarthy and J. Shamir, J. Opt. Soc. Am. **72**, 356 (1982); A. J. Dragt, *ibid.* **72**, 372 (1982); E. C. G. Sudarshan, R. Simon, and N. Mukunda, Opt. Acta **32**, 855 (1985).

[31]A. J. Dragt, *Lectures in Nonlinear Orbit Dynamics*, AIP Conf. Proc., Vol. 87, (1982); A. J. Dragt and E. Forest, in *Advances in Electronics and Electron Physics*, edited by P. W. Hawkes (Academic, New York, 1986), Vol. 57; A. J. Dragt, Nucl. Instrum. Methods **A258**, 339 (1987); A. J. Dragt, F. Neri, G. Rangarajan, D. R. Douglas, L. M. Healy, and R. D. Ryne, Ann. Rev. Nucl. Part. Sci. **38**, 455 (1988) and references therein.

[32]E. D. Courant and H. S. Snyder, Ann. Phys. **3**, 1 (1953); E. R. Caianiello and A. Turrin, CERN Report No. CERN/T/ERC/3, 1953 (unpublished).