Quantum theory of charged-particle beam optics

by

S A Khan

A THESIS IN PHYSICS

Submitted to the University of Madras in partial fulfillment of
the requirement for the degree of Doctor of Philosophy

January 1997

The Institute of Mathematical Sciences
C.I.T. Campus, Tharamani
Chennai (Madras), Tamilnadu - 600 113 INDIA
URL: http://www.imsc.ernet.in/

INSTITUTE OF MATHEMATICAL SCIENCES LIBRARY
CHENNAI - 600 113

CERTIFICATE

This is to certify that the Ph.D. thesis titled **Quantum theory of charged-particle beam optics** submitted to the University of Madras by S A KHAN is a record of bonafide research work done under my supervision. The research work presented in this thesis has not formed the basis for the award to the candidate of any Degree, Diploma, Associateship, Fellowship or other similar titles.

It is further certified that the thesis represents independent work by the candidate and collaboration when existed was necessitated by the nature and scope of problems dealt with.

R Jagannathan

Thesis Supervisor

January 1997

The Institute of Mathematical Sciences
C.I.T. Campus, Tharamani
Chennai (Madras), Tamilnadu - 600 113

THE INSTITUTE OF MATHEMATICAL SCIENCE
C.I.T. CAMPUS, MADRAS-600 113

Abstract

Charged-particle beam optics, or the theory of transport of charged-particle beams through electromagnetic systems, is traditionally dealt with using classical mechanics. Though the classical treatment has been very successful, in designing and working of numerous charged-particle optical devices, it is natural to look for a deeper understanding based on the quantum theory, since any system is quantum mechanical at the fundamental level. With this motivation, the quantum theory of charged-particle beam optics is being developed currently by Jagannathan *et al.*; this formalism is specifically adapted to treat the problems of beam optics. The present thesis is an elaboration of this new formalism of the quantum theory of charged-particle beam optics with illustrations of applications to several practically important systems. The essential content of the thesis can be summarized briefly as follows.

Quantum mechanics of the optics of charged-particle beams transported through an electromagnetic lens or other such optical systems is analyzed, at the level of single-particle dynamics, treating the electromagnetic fields as classical and disregarding the radiation aspects, using essentially an algebraic approach. The formalism is based on the basic equations of quantum mechanics appropriate to the situations. For situations when either there is no spin or spin can be treated as a spectator the scalar Klein-Gordon and Schrödinger equations are used as the basic equations for relativistic and nonrelativistic cases respectively. For spin-$\frac{1}{2}$ particles, a treatment based on the Dirac equation is presented taking fully into account the spinor character of the wavefunction. The underlying powerful algebraic machinery of the formalism makes it possible to do computations to any degree of accuracy in any situation from electron microscopy to accelerator optics. The power of the formalism is demonstrated by working out the examples which include the axially

symmetric magnetic round lens (of importance for electron microscopy and other micro-electron-beam device technologies) and the magnetic quadrupole lens (of importance for accelerator optics). It is found that the quantum theory at the scalar (spin-less) level gives rise to interesting small additional contributions to the classical paraxial and aberrating behaviours. These contributions are directly proportional to powers of the de Broglie wavelength. The Dirac theory further gives rise to spinor contributions which are also directly proportional to powers of the de Broglie wavelength.. Thus, it is clear that these quantum contributions are of significance only at very low energies; this explains the grand success of the classical theory so far. It is very interesting to note that the quantum correction terms arising from the Klein-Gordon theory and the scalar approximation of the Dirac theory do not coincide and have some small differences between them. The classical, or geometrical, charged-particle optics is obtained in the classical limit of the quantum theory as should be.

The formalism based on the Dirac theory is further applied to the study of the spin-dynamics of a Dirac particle with anomalous magnetic moment being transported through a magnetic optical element. This naturally leads to a unified treatment of both the orbital (the Lorentz and the Stern-Gerlach forces) and the spin (Thomas-Bargmann-Michel-Telegdi equation) motions. This is illustrated by computing, under the paraxial approximation, the transfer maps for the phase-space and spin components in the cases of normal and skew magnetic quadrupole lenses. The quantum mechanics of the concept of spin-splitter devices, proposed recently for achieving polarized beams, is also understood using our formalism.

An alternate approach to the quantum theory of charged-particle beam optics based on the Wigner phase-space distribution function is also presented briefly, restricting to the example of magnetic round lens treated under the paraxial approximation. The possibility of extension of such an approach to the Dirac, or the spinor case, is also noted.

The concluding section lists some interesting observations and points out a few directions for future research.

Publications leading to the Thesis

Parts of the research work leading to this thesis have been published, which include:

1. S. A. Khan and R. Jagannathan

   Theory of relativistic electron beam transport based on the Dirac equation

   Proceedings of the 3rd National Seminar on Physics and Technology of Particle Accelerators and their Applications (Nov. 1993, Calcutta), Ed. S. N. Chintalapudi (IUC-DAEF, Calcutta) 102-107.

2. S. A. Khan and R. Jagannathan

   Quantum mechanics of charged-particle beam optics: An operator approach

   Presented at the *JSPS-KEK International Spring School on High Energy Ion Beams – Novel Beam Techniques and their Applications* (March 1994, Japan), `Preprint: IMSc/94/11` (The Institute of Mathematical Sciences, Madras, Mar. 1994).

3. S. A. Khan and R. Jagannathan

   On the quantum mechanics of charged particle beam transport through magnetic lenses

   Phys. Rev. E **51**, 2510-2516 (1995).

4. R. Jagannathan and S. A. Khan

   Wigner functions in charged-particle optics

   Selected Topics in Mathematical Physics – Professor R. Vasudevan Memorial Volume, Eds. R. Sridhar, K. Srinivasa Rao and V. Lakshminarayanan (Allied Publishers, New Delhi, 1995) 308-321.

5. R. Jagannathan and S. A. Khan

Quantum theory of the optics of charged particles

Advances in Imaging and Electron Physics Vol.**97**, Ed. P. W. Hawkes (Academic Press, San Diego, 1996) 257-358.

6. M. Conte, R. Jagannathan, S. A. Khan and M. Pusterla

Beam optics of the Dirac particle with anomalous magnetic moment

Particle Accelerators **56**, 99-126 (1996).

7. S. A. Khan

Transport of Dirac-particle beams through magnetic quadrupoles

`Preprint: IMSc/96/33` (The Institute of Mathematical Sciences, Madras, Dec. 1996).

Acknowledgements

I am indebted to my thesis supervisor Prof. R. Jagannathan not only for introducing me to the fascinating field of charged-particle beam optics, but also for patience he has shown in guiding me through these four years. I shall always envy the thoroughness with which he looks at any problem. He has always given his full attention and has been of great support in many ways.

I have had the pleasure of many discussions with Prof. R. Simon, on several aspects of the analogy between electron optics and photon optics. I am grateful to him for continuous encouragement.

I am thankful to Prof. R. Ramachandran, Prof. G. Rajasekaran and Prof. K. Srinivasa Rao for kind encouragement.

I am grateful to the Office & Library Staff, for being very helpful at various stages of my work at Matscience. I extend my thanks to all of them.

Finally, I wish to thank all my friends here for the kind of unadulterated fun that we had in Matscience.

INSTITUTE OF MATHEMATICAL SCIENCES LIBRARY
* CHENNAI - 600 113 *

INSTITUTE OF MATHEMATICAL SCIENCES * CHENNAI *

Contents

0 Introduction 1

0.1 Thesis Layout 4

1 Review of the classical theory 7

1.1 Classical mechanics of charged-particle beams in electromagnetic fields 7

1.2 Lie algebraic methods 14

2 Scalar quantum theory of charged-particle beam optics 27

2.1 Formalism 27

2.2 Applications 46

2.2.1 Free Propagation: Diffraction 46

2.2.2 Axially Symmetric Magnetic Lens: Electron Optical Imaging . 49

2.2.3 Magnetic Quadrupole Lens 79

2.2.4 Axially Symmetric Electrostatic Lens 81

2.2.5 Electrostatic Quadrupole Lens 83

3 Spinor theory of charged-particle beam optics 84

3.1 Formalism 84

3.2 Applications 93

3.2.1 Free Propagation : Diffraction 93

3.2.2 Axially Symmetric Magnetic Lens 95

3.2.3 Magnetic Quadrupole Lens 98

4 **Spin dynamics of the Dirac-particle beam** **99**

4.1 Introduction 99

4.2 Formalism 101

4.3 Applications 112

4.3.1 Normal Magnetic Quadrupole 112

4.3.2 Skew Magnetic Quadrupole 117

4.4 Stern-Gerlach Force 120

5 **Phase-space formalism of the quantum theory of charged-particle beam optics** **125**

6 **Concluding remarks** **135**

Appendices **143**

A The Magnus Formula 143

B The Feshbach-Villars Form of the Klein-Gordon Equation 146

C The Foldy-Wouthuysen Representation of the Dirac Equation 149

D Green's Function for the Nonrelativistic Free Particle 156

E Matrix element of the Rotation Operator 158

F Green's Function for a Time-Dependent Quadratic Hamiltonian 159

Bibliography **164**

Chapter 0

Introduction

Charged-particle beam optics, or the theory of transport of charged-particle beams through electromagnetic systems, is traditionally dealt with using classical mechanics. This is the case in electron and ion optics, electron microscopy, accelerator physics etc. (see, e.g., [1, 2, 3]). Though the classical treatment has been very successful, in the designing and working of numerous optical components, it is natural to look for a prescription based on quantum theory, since any physical system is quantum mechanical at the fundamental level. Such a prescription can be believed to provide a deeper understanding of the working of charged-particle beam devices.

During 1930's Glaser pioneered the development of the quantum theory of image formation in electron microscopy on the basis of the nonrelativistic Schrödinger equation (see Glaser's classic work [4]). Details of Glaser's theory and consequent developments in electron microscopy are available, with extensive bibliography and historical notes, in the recently published third volume of the three-volume encyclopedic text book of Hawkes and Kasper [5].

It is curious to note that the use of the Dirac equation, the proper basic equation of the electron, has not drawn adequately the attention of the researchers in electron microscopy and other micro-electron-beam devices. After some preliminary studies by Rubinowicz (1934), Durand (1953), and Phan-Van-Loc (1953) on the use of the Dirac equation in electron optics (mostly the study of electron diffraction, see [5] for

INSTITUTE OF MATHEMATICAL SCIENCES LIBRARY

detailed bibliography), it was only in the last decade that Ferwerda *et al.* (1986) first reopened the question of using the spinor wavefunction for understanding electron optical images. Essentially, Ferwerda *et al.* [6, 7] found after a thorough analysis that the use of the scalar Klein-Gordon wavefunction in electron microscopy could be vindicated since a scalar approximation of the Dirac spinor theory would be justifiable under the conditions obtaining in present day electron microscopes.

Subsequently, the development of the spinor electron optics is being pursued by Jagannathan *et al.* [8], Jagannathan [9, 10], and Khan and Jagannathan [11, 12] mainly due to a desire to understand how the Dirac equation, the equation for electrons, explains electron 'optics'. Of course, there is also the hope that any better understanding of the way the scalar theory becomes such an excellent approximation of the spinor theory in electron microscopy may eventually be of some practical use in certain situations. For spin-0 particles, or when the spin can be treated as a spectator, scalar electron wave optics has also been developed based on the Klein-Gordon equation and the Schrödinger equation corresponding to the relativistic and nonrelativistic situations respectively [13, 14].

The formalism of Jagannathan *et al.* was the first one, to derive the focusing theory for electron lenses, in particular for magnetic and electrostatic round lenses and quadrupole lenses, from the Dirac equation. The formalism of Jagannathan *et al.* gives a rigorous recipe based on quantum mechanics to calculate the lens properties, including aberrations, up to any degree of accuracy through a systematic series method. The present thesis is an elaboration of this quantum mechanical formalism and deals with applications to several physical systems such as free propagation (diffraction), axially symmetric magnetic and electrostatic lenses and magnetic and electrostatic quadrupole lenses. Throughout the thesis, electromagnetic fields are treated as classical and radiation effects are neglected. Further, the treatment is at the level of single-particle dynamics.

The traditional geometrical charged-particle optics is obtained in the classical limit of the quantum theory as should be. It is found that there are interesting additional small contributions to the classical aberrations, arising from quantum mechanics even at the scalar level. Of course, in the Dirac theory the spinor nature of the wavefunction modifies further, though only minutely, the various optical characteristics of the system. In the classical limit the algebraic approach of our theory tends to the Lie algebraic treatment of classical charged-particle beam optics pioneered by Dragt *et al.* ([15]-[19]) (see also Forest and Hirata [20] and Forest *et al.* [21]).

Traditionally the main framework for studying the spin dynamics and beam polarization has been the well-known quasiclassical Thomas-Bargmann-Michel-Telegdi (Thomas-BMT) equation (see,*e.g.*, [22]). Application of the spinor beam optical formalism has been shown to lead [23] to a fully quantum mechanical understanding of the dynamics of a spin-$\frac{1}{2}$ particle with anomalous magnetic moment, including the spin evolution, at the level of single-particle dynamics. The general theory, presented here for any magnetic optical element with straight axis, describes the quantum mechanics of the orbital dynamics, the Stern-Gerlach kicks and the Thomas-BMT spin evolution, up to the paraxial approximation. To illustrate the general theory, the first order transfer maps for the spin components and the transverse phase-space, including the transverse Stern-Gerlach kicks, are computed for normal magnetic quadrupole lens (see [23]) and skew magnetic quadrupole lens (see [24]). The longitudinal Stern-Gerlach kick in a general inhomogeneous magnetic field is also discussed briefly. Stern-Gerlach kicks are the basic mechanisms in the spin-splitter devices proposed recently as alternative methods for getting polarized beams (see [25] and references therein).

A phase-space formalism of the quantum mechanics of charged-particle beam optics using the Wigner distribution function is also presented based on some preliminary work in this direction by us [26]. Of course, there exist other works on this topic [27] even going beyond the paraxial approximation; our work deals only with the paraxial situation.

0.1 Thesis Layout

The thesis is broadly divided into six chapters with several appendices to supplement the calculational details presented in the main body of the thesis. The contents of the various chapters are as follows.

The present zeroth chapter, or the introductory chapter, briefly covers the historical development of the quantum theory of charged-particle beam optics, and the motivation for the research which has culminated into the present thesis.

The first chapter is devoted to the review of the classical theory of charged-particle beam optics, closely following the Lie algebraic approach pioneered by Dragt *et al.* This chapter starts with the standard relativistic classical Lagrangian and Hamiltonian which are cast into beam-optical forms to study the evolution of the beam along the optic axis.

In the second chapter the scalar quantum theory of charged-particle beam transport through an electromagnetic lens system with a straight optics axis at the level of single-particle dynamics is developed starting with the basic equations of quantum mechanics appropriate to the situation under study. For situations when either there is no spin or spin can be disregarded the Klein-Gordon and Schrödinger equations are used as the basic equations for relativistic and nonrelativistic cases respectively.

In the third chapter a spinor theory of charged-particle beam transport through an electromagnetic lens system is developed appropriately based on the Dirac equation, the basic equation for the spin-$\frac{1}{2}$ particles, taking fully into account the spinor character of the wavefunction. The formalism is applied to the study of electron-optical imaging which involves the magnetic round lens. Magnetic quadrupole and electrostatic round and quadrupole lenses are also studied briefly.

The fourth chapter develops the quantum theoretic framework for studying the spin dynamics and beam polarization in accelerator physics. Starting with the standard Dirac-Pauli equation for a spin-$\frac{1}{2}$ particle with anomalous magnetic moment it is shown how to obtain a representation in which the effective 'accelerator optical' Hamiltonian accounts, in a unified way, for both the orbital (the Lorentz and the Stern-Gerlach forces) and the spin (the Thomas-BMT equation) motions. The general theory developed for any magnetic element with straight optic axis and up to the lowest order (paraxial) approximation is illustrated by computing the transfer maps for phase-space and spin components in the case of magnetic quadrupole lenses. The quantum mechanics of Stern-Gerlach kicks is also discussed.

The fifth chapter is devoted to the application of the Wigner phase-space distribution for studying the quantum mechanics of charged-particle beam transport through a magnetic optical system. Such a study provides a natural link between the classical and the quantum descriptions. In this context, the relation between the transformation of the Wigner function of a charged particle optical system, corresponding to the associated scalar wavefunction, and the transformation of the classical phase-space of the system is studied. As an example, the magnetic round lens is studied in the paraxial approximation. The focusing action of the lens and the expressions of focal length, image rotation, etc., are understood using this formalism. The chapter concludes with comments on the possibility of extension of the phase-space formalism to the study of aberrating systems and the Dirac, or spinor, charged-particle optics.

The sixth chapter is a collection of some concluding remarks and interesting observations.

The thesis ends with several appendices, supplementing the calculational details presented in the main text, which include: the Magnus formula, the Feshbach-Villars form of the Klein-Gordon equation, the Foldy-Wouthuysen representation of the Dirac equation, and Green's function for a system with time-dependent Hamiltonian quadratic in position and momentum.

Lastly, there is the bibliography.

Chapter 1

Review of the classical theory

1.1 Classical mechanics of charged-particle beams in electromagnetic fields

The trajectories of a charged-particle of rest mass m_0 and charge q moving in presence of an electric field $\boldsymbol{E}(\boldsymbol{r}, t)$ and a magnetic field $\boldsymbol{B}(\boldsymbol{r}, t)$ are completely described by the Lorentz force equation

$$\boldsymbol{F} = \frac{d}{dt}\left(\frac{m_0 \boldsymbol{v}}{\sqrt{1 - \frac{v^2}{c^2}}}\right) = q\left\{\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right\} . \tag{1.1}$$

Let the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ be derived from the electric scalar potential $\phi(\boldsymbol{r}, t)$ and the magnetic vector potential $\boldsymbol{A}(\boldsymbol{r}, t)$:

$$\boldsymbol{E} = -\frac{\partial}{\partial t}\boldsymbol{A} - \nabla\phi\,, \quad \boldsymbol{B} = \nabla \times \boldsymbol{A}\,. \tag{1.2}$$

It is to be noted that even for systems with very simple geometries the potentials can be expressed only through an infinite series.

The Lorentz force equation follows from the Lagrangian

$$L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = -m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} + q\left(\boldsymbol{v} \cdot \boldsymbol{A}(\boldsymbol{r}, t) - \phi(\boldsymbol{r}, t)\right) \tag{1.3}$$

according to the Euler-Lagrange equations of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \boldsymbol{v}}\right) - \frac{\partial L}{\partial \boldsymbol{r}} = 0 \tag{1.4}$$

derived from Hamilton's variational principle. The Hamiltonian is:

$$H(\boldsymbol{r},\boldsymbol{p},t) = \frac{\partial L}{\partial \boldsymbol{v}} \cdot \boldsymbol{v} - L(\boldsymbol{r},\dot{\boldsymbol{r}},t) = \sqrt{m_0^2 c^4 + |\pi|^2 c^2} + q\phi(\boldsymbol{r},t), \tag{1.5}$$

where $\boldsymbol{\pi}$, the kinetic momentum, and $\boldsymbol{p} = \frac{\partial L}{\partial \boldsymbol{v}}$, the canonical momentum, are related to each other as

$$\pi = \boldsymbol{p} - q\boldsymbol{A}. \tag{1.6}$$

The corresponding Hamilton's equations of motion are:

$$\frac{d\boldsymbol{r}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}}, \qquad \frac{d\boldsymbol{p}}{dt} = -\frac{\partial H}{\partial \boldsymbol{r}}. \tag{1.7}$$

Let us now write the Hamilton's equations in the language of Poisson brackets. Given two functions $f(\boldsymbol{r},\boldsymbol{p})$ and $g(\boldsymbol{r},\boldsymbol{p})$ in the $(\boldsymbol{r},\boldsymbol{p})$-phase-space of the system, their Poisson bracket $\{\ ,\ \}$ is defined by

$$\{f,g\} = \sum_\alpha \left(\frac{\partial f}{\partial r_\alpha}\frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha}\frac{\partial g}{\partial r_\alpha} \right), \tag{1.8}$$

where $\sum_\alpha$ stands for sum over all components of $\boldsymbol{r}$ and $\boldsymbol{p}$. Hamilton's equations can be now alternately written as

$$\frac{d\boldsymbol{r}}{dt} = \{\boldsymbol{r}, H\}, \qquad \frac{d\boldsymbol{p}}{dt} = \{\boldsymbol{p}, H\}. \tag{1.9}$$

In general, for any observable $f(\boldsymbol{r},\boldsymbol{p},t)$ of the system the equation of motion is

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, H\}. \tag{1.10}$$

Hamilton's equations of motion are in general coupled and time-dependent making their solutions difficult even in relatively simple situations. One often has to resort to suitable approximation techniques or eventually use numerical methods.

The situation described above is very general. For a charged-particle beam device the problem is somewhat simplified by invoking the geometry of the system under study and suitable approximations to the extent permissible. In many devices the electromagnetic fields are static or can be reasonably assumed to be static. This makes the potentials to be time-independent. In many such devices one can further ignore the times of flights which are either negligible or not of interest as the emphasis is on the profiles of the trajectories. For such situations the first step lies in transforming the above Lagrangian $L\left(\boldsymbol{r}, \dot{\boldsymbol{r}}, t\right)$ to a time-independent one. This is done by eliminating t in preference to a variable say s which is the arc-length along the design trajectory of the particle whose motion is being studied. For systems with straight optic axis s will be the coordinate along the optics axis. The present study mainly deals with systems with straight optic axis.

We shall choose the Cartesian coordinate system, assume the optic axis to be situated along the z-direction and introduce the notation $\boldsymbol{r}_\perp = (x, y)$ where x and y are the off-axis coordinates. In general, the subscript $\perp$ will stand for the off-axis quantities which could be the x and y components of the momentum, or the magnetic vector potential, for instance, and for any transverse vector $\boldsymbol{V}_\perp$, $V_\perp^2 = V_x^2 + V_y^2$. And, for any quantity the prime ($'$) will stand for differentiation with respect to z, the coordinate along the optic axis.

Now, using the well-known Maupertuis principle of classical mechanics one can cast the time-dependent Lagrangian L into a new time-independent beam-optical Lagrangian:

$$\mathcal{L}\left(\boldsymbol{r}_\perp, \boldsymbol{r}'_\perp, z\right) = \frac{\partial L}{\partial \boldsymbol{v}} \cdot \boldsymbol{r}' = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} \sqrt{1 + {\boldsymbol{r}'_\perp}^2} + q\left(\boldsymbol{A}_\perp \cdot \boldsymbol{r}'_\perp + A_z\right) . \tag{1.11}$$

The corresponding Euler-Lagrange equations are:

$$\frac{d}{dz}\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}'_\perp}\right) - \frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_\perp} = 0 . \tag{1.12}$$

For beam propagation $|\boldsymbol{r}'_\perp| < 1$ and in the paraxial case $|\boldsymbol{r}'_\perp| \ll 1$. Having obtained the beam-optical Lagrangian it is natural to derive the beam-optical Hamiltonian:

one can show that

$$\begin{aligned} \mathcal{H}\left(\boldsymbol{r}_{\perp}, \boldsymbol{p}_{\perp}, z\right) &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_{\perp}'} \cdot \boldsymbol{r}_{\perp}' - \mathcal{L}\left(\boldsymbol{r}_{\perp}, \boldsymbol{r}_{\perp}', z\right) \\ &= -\frac{1}{c}\sqrt{(E - q\phi)^2 - m_0^2 c^4 - c^2|\boldsymbol{\pi}_{\perp}|^2} - qA_z \,, \end{aligned} \tag{1.13}$$

where E is the total energy of the particle (including the rest energy); it must be noted that for a system with stationary field the total energy is conserved. The Hamiltonian $\mathcal{H}$ in (1.13) just corresponds to $-p_z$, negative of the z-component of canonical momentum, as seen by solving the relativistic expression for energy $E = \sqrt{m_0^2 c^4 + c^2|\boldsymbol{\pi}|^2} + q\phi$. Later on, we shall see in quantum theory that the corresponding beam-optical Hamiltonian operator $\hat{\mathcal{H}}$ would correspond to $-\hat{p}_z = \mathrm{i}\hbar\frac{\partial}{\partial z}$ so that the z-evolution equation becomes $\mathrm{i}\hbar\frac{\partial}{\partial z}\psi(\boldsymbol{r}_{\perp}, z) = \hat{\mathcal{H}}\psi(\boldsymbol{r}_{\perp}, z)$ in analogy with the Schrödinger equation for temporal evolution, namely, $\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r}, t) = \hat{H}\psi(\boldsymbol{r}, t)$. If the kinetic (= canonical) momentum of the particle in the field-free regions outside the system, where $\phi = 0$ and $\boldsymbol{A} = 0$, is denoted by p_0 then $E = \sqrt{m_0^2 c^4 + c^2 p_0^2}$; p_0 is called the design momentum and for beam propagation $p_z \approx p_0 \gg |\boldsymbol{p}_{\perp}|$. The Hamilton's equations for the transverse phase-space coordinates are:

$$\frac{d\boldsymbol{r}_{\perp}}{dz} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_{\perp}}\,, \qquad \frac{d\boldsymbol{p}_{\perp}}{dz} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{r}_{\perp}}\,. \tag{1.14}$$

Or, equivalently, written in the Poisson bracket form

$$\frac{d\boldsymbol{r}_{\perp}}{dz} = \left\{\boldsymbol{r}_{\perp}, \mathcal{H}\right\}, \qquad \frac{d\boldsymbol{p}_{\perp}}{dz} = \left\{\boldsymbol{p}_{\perp}, \mathcal{H}\right\}, \tag{1.15}$$

where, now,

$$\{f, g\} = \left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial p_x} - \frac{\partial f}{\partial p_x}\frac{\partial g}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial p_y} - \frac{\partial f}{\partial p_y}\frac{\partial g}{\partial y}\right). \tag{1.16}$$

The beam-optical Lagrangian and the beam-optical Hamiltonian are the starting points for the study of charged-particle beam devices.

To understand the optical behaviour of any system the corresponding Lagrangian $\mathcal{L}$ is expanded in a power series in the off-axis coordinates and their z-derivatives, i.e., one writes

$$\mathcal{L}\left(\boldsymbol{r}_{\perp}, \boldsymbol{r}'_{\perp}, z\right) = \sum_{i=0}^{\infty} \mathcal{L}_{(i)}\left(\boldsymbol{r}_{\perp}, \boldsymbol{r}'_{\perp}, z\right) \tag{1.17}$$

where $\mathcal{L}_{(i)}\left(\boldsymbol{r}_{\perp}, \boldsymbol{r}'_{\perp}, z\right)$ is a homogeneous polynomial of degree i in $\boldsymbol{r}_{\perp}$ and $\boldsymbol{r}'_{\perp}$. By retaining only terms up to $i = 2$ one gets the paraxial approximation. Then, the corresponding differential equations are linear and the system is easily described in terms of matrix theory. Some properties of the system can be deduced from the symmetry of the system reflected in the form of $\mathcal{L}_{(i)}\left(\boldsymbol{r}_{\perp}, \boldsymbol{r}'_{\perp}, z\right)$. For instance, when a system possesses axial symmetry (to be discussed in detail later) all odd order terms in (1.17), namely $\mathcal{L}_{(2i+1)}$, vanish. For such a system the first significant contribution comes from $\mathcal{L}_{(2)}$ which governs the paraxial behaviour.

The non-paraxial contributions are treated perturbatively. In these studies there are two well-known methods adopted: namely, the eikonal or characteristic function method due to Glaser and Sturrock and the trajectory method due to Scherzer. Details and references on these are available in [1].

Except in Chapter V, we shall be mostly using the Hamiltonian formalism. So we expand the beam-optical Hamiltonian in (1.13) in a power series in the off-axis coordinates and canonical momenta as

$$\mathcal{H}\left(\boldsymbol{r}_{\perp}, \boldsymbol{p}_{\perp}, z\right) = \sum_{i=0}^{\infty} \mathcal{H}_{(i)}\left(\boldsymbol{r}_{\perp}, \boldsymbol{p}_{\perp}, z\right) , \tag{1.18}$$

where $\mathcal{H}_{(i)}\left(\boldsymbol{r}_{\perp}, \boldsymbol{p}_{\perp}, z\right)$ is a homogeneous polynomial of degree i in $\boldsymbol{r}_{\perp}$ and $\boldsymbol{p}_{\perp}$. Note that $1/p_0$, where p_0 is the design momentum, will serve as the expansion parameter as is easily seen by expanding the Hamiltonian in (1.13) corresponding to the pure magnetic system for which $\phi = 0$. In the expansion (1.18) the zeroth order term $\mathcal{H}_{(0)}$ is a constant and hence can be ignored. the first-order term $\mathcal{H}_{(1)}$ is linear and results in translations. Hence $\mathcal{H}_{(1)}$ can also be set aside. So, the first physically significant term is $\mathcal{H}_{(2)}$. By retaining terms up to $i = 2$ one gets the paraxial approximation;

i.e., the corresponding differential equations are linear and the system is described in terms of matrix theory. In general, we define the paraxial Hamiltonian as

$$\mathcal{H}_p = \mathcal{H}_{(0)} + \mathcal{H}_{(1)} + \mathcal{H}_{(2)} . \tag{1.19}$$

The remaining terms in (1.18), giving rise to the nonlinear (aberrating) behaviour, are denoted by the aberrating Hamiltonian

$$\mathcal{H}_a = \mathcal{H}_{(3)} + \mathcal{H}_{(4)} + \mathcal{H}_{(5)} + \cdots . \tag{1.20}$$

Some properties of the system can be deduced from the symmetry of the system reflected in the form of $\mathcal{H}_{(i)}\left(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp, z\right)$. For instance, in the case of a system with axial symmetry $\mathcal{H}_{(2i+1)} \equiv 0$. For such a system the first significant contribution comes from $\mathcal{H}_{(2)}$ which governs the paraxial behaviour. In general $\mathcal{H}_p$ will be of the form

$$\begin{aligned} \mathcal{H}_p &= \text{constant} + \text{linear terms} \\ &\quad + a p_\perp^2 + b r_\perp^2 + c L_z + d\, \boldsymbol{r}_\perp \cdot \boldsymbol{p}_\perp + \cdots , \end{aligned} \tag{1.21}$$

where the z-dependent constants $a, b, c, d, \ldots$, etc., characterize the system and $L_z = x p_y - y p_x$ is the z-component of angular momentum. For an axially symmetric system the Hamiltonian will contain only terms which are invariant under rotation around the optic axis; such terms will have vanishing Poisson brackets with L_z so that the Poisson bracket of the entire Hamiltonian with L_z is zero.

To illustrate the forgoing discussion on Hamiltonians, let us consider the case of the axially symmetric magnetic lens, or the round magnetic lens, which is characterized by the potentials

$$\phi(\boldsymbol{r}) = 0 \tag{1.22}$$

$$\boldsymbol{A} = \left(-\frac{y}{2}\Pi(\boldsymbol{r}_\perp, z)\,,\ \frac{x}{2}\Pi(\boldsymbol{r}_\perp, z)\,,\ 0\right) , \tag{1.23}$$

with

$$\begin{aligned}
\Pi(\boldsymbol{r}_\perp, z) &= \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!}\left(-\frac{r_\perp^2}{4}\right)^n B^{(2n)}(z) \\
&= B(z) - \frac{1}{8} r_\perp^2 B''(z) + \frac{1}{192} r_\perp^4 B''''(z) - \ldots ,
\end{aligned} \tag{1.24}$$

where $B^0(z) = B(z)$, $B'(z) = \frac{dB(z)}{dz}$, $B''(z) = \frac{d^2B(z)}{dz^2}$, $B'''(z) = \frac{d^3B(z)}{dz^3}$, $B''''(z) = \frac{d^4B(z)}{dz^4}, \ldots$, and, in general, $B^{(2n)}(z) = \frac{d^{(2n)}B(z)}{dz^{(2n)}}$ (for more details, see [1]). The corresponding magnetic field, rotationally symmetric with respect to the optic axis of the system (z-axis), is

$$\begin{aligned}
\boldsymbol{B}_\perp &= -\frac{1}{2}\left(B'(z) - \frac{1}{8} r_\perp^2 B'''(z) + \ldots\right) \boldsymbol{r}_\perp \\
B_z &= B(z) - \frac{1}{4} r_\perp^2 B''(z) + \frac{1}{64} r_\perp^4 B''''(z) - \ldots ,
\end{aligned} \tag{1.25}$$

as given by $\boldsymbol{B} = \nabla \times \boldsymbol{A}$; it may be noted that the function $B(z)$ completely characterizes the field. In general, we shall use the notation $f'(z) = \frac{df(z)}{dz}$, $f''(z) = \frac{d^2f(z)}{dz^2}$, ... , for any $f(z)$.

The corresponding beam-optical Hamiltonian for the round magnetic lens is given by

$$\mathcal{H} = \mathcal{H}_{(0)} + \mathcal{H}_{(2)} + \mathcal{H}_{(4)} + \cdots , \tag{1.26}$$

$$\mathcal{H}_{(0)} = -p_0 , \tag{1.27}$$

$$\mathcal{H}_{(2)} = \frac{1}{2p_0} p_\perp^2 + \frac{p_0}{2} \alpha^2 r_\perp^2 - \alpha L_z , \tag{1.28}$$

$$\begin{aligned}
\mathcal{H}_{(4)} &= \frac{1}{8p_0^3} p_\perp^4 - \frac{1}{2p_0^2} \alpha p_\perp^2 L_z - \frac{1}{2p_0} \alpha^2 \left(\boldsymbol{r}_\perp \cdot \boldsymbol{p}_\perp\right)^2 \\
&\quad + \frac{3}{4p_0} \alpha^2 \left(r_\perp^2 p_\perp^2\right) + \frac{1}{8}\left(\alpha'' - 4\alpha^3\right) L_z r_\perp^2 + \frac{p_0}{8}\left(\alpha^4 - \alpha\alpha''\right) r_\perp^4 , \\
&\qquad \text{with } \alpha = \frac{qB(z)}{2p_0} .
\end{aligned} \tag{1.29}$$

From equations (1.26) to (1.29) we note that $\mathcal{H}_p$ is the sum of $-p_0$ and a homogeneous quadratic polynomial in $(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp)$ and the leading order contribution to $\mathcal{H}_a$ comes from $\mathcal{H}_{(4)}$.

1.2 Lie algebraic methods

Elegant Lie algebraic techniques have been developed, in the pioneering works of Dragt *et al.*, to analyse systematically the behaviour of beam optical devices treated as classical systems. The present thesis contains an extension of these techniques to the analysis of charged-particle beam optical devices treated as quantum mechanical systems. In this section a brief exposition of the Lie algebraic methods, their use, and some examples relevant to the thesis, will be given (for detailed accounts see Dragt *et al.* ([15]-[19]), Forest and Hirata [20] and Forest *et al.* [21]).

Let $\boldsymbol{w} = \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}$, a four-component column vector. Or, saying more explicitly, the row vector $\boldsymbol{w}^T = (\boldsymbol{r}_\perp, \boldsymbol{p}_\perp/p_0)$ represents $(x, y, p_x/p_0, p_y/p_0)$. Let $\boldsymbol{w}_{\rm in}$ and $\boldsymbol{w}_{\rm out}$ denote the initial and final values of $\boldsymbol{w}$ at the points $z_{\rm in}$ and $z_{\rm out}$ denoting the coordinates of the optical system on its optic axis at the beginning and at the end respectively. The basic problem on hand is: given the beam-optical Hamiltonian $\mathcal{H}$ of the system what is the relation between $\boldsymbol{w}_{\rm out}$ and $\boldsymbol{w}_{\rm in}$, or vice versa ?. We seek an optical transfer map of the form

$$\boldsymbol{w}_{\rm out} = (\mathcal{M}\boldsymbol{w})_{\rm in}\,, \tag{1.30}$$

for the given system. As is well-known, for a linear system, characterized by the paraxial Hamiltonian $\mathcal{H}_p$, such a map is a 4×4 symplectic matrix called the transfer matrix. The Lie algebraic approach generalizes the matrix methods, which describe the paraxial approximation, to operator methods suitable for describing linear and nonlinear (aberrating) systems.

To get an idea of the Lie techniques let us begin by defining a few Lie algebraic tools. Let $\mathcal{F}(\boldsymbol{w})$ $(= \mathcal{F}(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp))$ be some specified function. The Lie derivative operator associated with $\mathcal{F}(\boldsymbol{w})$ is defined by

$$: \mathcal{F} : = \sum_{\alpha} \left(\frac{\partial \mathcal{F}}{\partial r_\alpha} \frac{\partial}{\partial p_\alpha} - \frac{\partial \mathcal{F}}{\partial p_\alpha} \frac{\partial}{\partial r_\alpha} \right) \tag{1.31}$$

and its action on some function $g(\boldsymbol{w})$ $(= g(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp))$ results in the Poisson bracket

$$: \mathcal{F} : g \;=\; \{\mathcal{F}, g\} \,. \tag{1.32}$$

It is straightforward to check that

$$: \mathcal{F} :^2 g =: \mathcal{F} : : \mathcal{F} : g = \{\mathcal{F}, \{\mathcal{F}, g\}\} \,. \tag{1.33}$$

Thus, we can define higher powers of $: \mathcal{F} :$ through the Poisson brackets, noting that the zeroth power of $: \mathcal{F} :$ is the identity operator ($\mathcal{I}$)

$$: \mathcal{F} :^0 g = g \,. \tag{1.34}$$

Having defined the powers of the Lie operator $: \mathcal{F} :$ we can define power series and consequently functions of $: \mathcal{F} :$. In the present context we shall be interested in the exponentiation of the Lie operator

$$\exp\left(: \mathcal{F} :\right) = \sum_{n=0}^{\infty} \frac{1}{n!} : \mathcal{F} :^n \tag{1.35}$$

and its action on the function g is

$$\exp\left(: \mathcal{F} :\right) g = g + \{\mathcal{F}, g\} + \frac{1}{2!} \{\mathcal{F}, \{\mathcal{F}, g\}\} + \cdots \,. \tag{1.36}$$

The operator $\exp\left(: \mathcal{F} :\right)$ in (1.36)) is called the Lie transformation associated with $\mathcal{F}$. The inverse Lie transformation $e^{(:-\mathcal{F}:)}$ is such that

$$e^{(:\mathcal{F}:)} e^{(:-\mathcal{F}:)} = e^{(:-\mathcal{F}:)} e^{(:\mathcal{F}:)} = \mathcal{I} \,. \tag{1.37}$$

Let us list some properties of the Lie transformation. With g and h as some functions and a and b as some constants, we have the following:

Linearity:

$$e^{(:\mathcal{F}:)}\left(ag+bh\right)=ae^{(:\mathcal{F}:)}g+be^{(:\mathcal{F}:)}h \tag{1.38}$$

Product preservation:

$$e^{(:\mathcal{F}:)}\left(gh\right)=\left(e^{(:\mathcal{F}:)}g\right)\left(e^{(:\mathcal{F}:)}h\right) \tag{1.39}$$

Poisson bracket preservation:

$$e^{(:\mathcal{F}:)}\left\{g\,,h\right\}=\left\{e^{(:\mathcal{F}:)}g,e^{(:\mathcal{F}:)}h\right\} \tag{1.40}$$

Composition:

$$e^{(:\mathcal{F}:)}g\left(\boldsymbol{w}\right)=g\left(e^{(:\mathcal{F}:)}\boldsymbol{w}\right)=g\left(e^{(:\mathcal{F}:)}\boldsymbol{r}_\perp,e^{(:\mathcal{F}:)}\boldsymbol{p}_\perp\right). \tag{1.41}$$

These basic properties can be used to derive many results. We shall just state a few of those which will be used by us. Let $\mathcal{M}$ be a general map of the form, say, $e^{(:\mathcal{F}:)}$. Then, for any $\mathcal{G}(\boldsymbol{w})$

$$\mathcal{M}e^{(:\mathcal{G}(\boldsymbol{w}):)}\mathcal{M}^{-1}=e^{(:\mathcal{F}:)}e^{(:\mathcal{G}:)}e^{(:-\mathcal{F}:)}=e^{\left(:e^{(:\mathcal{F}:)}\mathcal{G}:\right)}=e^{(:\mathcal{G}(\mathcal{M}\boldsymbol{w}):)}. \tag{1.42}$$

The other property to be mentioned is that the commutator (denoted by $[\ ,\ ]$) of two Lie operators is the Lie operator associated with the function obtained by taking the Poisson bracket of the corresponding functions:

$$[:\mathcal{F}:,:\mathcal{G}:]=:\mathcal{F}::\mathcal{G}:-:\mathcal{G}::\mathcal{F}:\equiv:\{\mathcal{F}\,,\mathcal{G}\}:. \tag{1.43}$$

Also note that if $\mathcal{H}_{(m)}$ and $\mathcal{H}_{(n)}$ are two homogeneous polynomials in $(\boldsymbol{r}_\perp,\boldsymbol{p}_\perp)$ of degree m and n, respectively, then,

$$\deg\left\{\mathcal{H}_{(m)}\,,\mathcal{H}_{(n)}\right\}=m+n-2\,. \tag{1.44}$$

We shall be making use of the above relations often.

Now, we can write the Hamiltonian equation of motion in the language of Lie transformations. For any observable f without any explicit z-dependence the Hamilton's equation is

$$\frac{df}{dz} =: -\mathcal{H} : f \,. \tag{1.45}$$

as seen from (1.15). Then, it follows that the transfer map $\mathcal{M}$ relating $w_{\rm out}$ at $z_{\rm out}$ to $w_{\rm in}$ at $z_{\rm in}$ is obtained by formal integration of this z-evolution equation (1.45): the result is

$$\begin{aligned}
\mathcal{M} &= \wp\left\{\exp\left(-\int_{z_{\rm in}}^{z_{\rm out}} dz : \mathcal{H}(z) :\right)\right\} \\
&= \exp\left\{-\int_{z_{\rm in}}^{z_{\rm out}} dz_1 : \mathcal{H}(z_1) : \right. \\
&\quad -\frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz_1 \int_{z_{\rm in}}^{z_1} dz_2 \left[: \mathcal{H}(z_1) : , : \mathcal{H}(z_2) :\right] \\
&\quad -\frac{1}{6}\int_{z_{\rm in}}^{z_{\rm out}} dz_1 \int_{z_{\rm in}}^{z_1} dz_2 \int_{z_{\rm in}}^{z_2} dz_3 \left(\left[\left[: \mathcal{H}(z_1) : , : \mathcal{H}(z_2) :\right] , : \mathcal{H}(z_3) :\right]\right. \\
&\qquad \left.\left. + \left[\left[: \mathcal{H}(z_3) : , : \mathcal{H}(z_2) :\right] , : \mathcal{H}(z_1) :\right]\right) + \cdots\right\} \\
&= \exp\left\{-\int_{z_{\rm in}}^{z_{\rm out}} dz_1 : \mathcal{H}(z_1) : \right. \\
&\quad -\frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz_1 \int_{z_{\rm in}}^{z_1} dz_2 : \left\{\mathcal{H}(z_1) , \mathcal{H}(z_2)\right\} : \\
&\quad -\frac{1}{6}\int_{z_{\rm in}}^{z_{\rm out}} dz_1 \int_{z_{\rm in}}^{z_1} dz_2 \int_{z_{\rm in}}^{z_2} dz_3 \left(: \left\{\left\{\mathcal{H}(z_1) , \mathcal{H}(z_2)\right\} , \mathcal{H}(z_3)\right\} :\right. \\
&\qquad \left.\left. + : \left\{\left\{\mathcal{H}(z_3) , \mathcal{H}(z_2)\right\} , \mathcal{H}(z_1)\right\} :\right) + \cdots\right\} ,
\end{aligned} \tag{1.46}$$

where $\wp$ denotes the path-ordered exponential. The explicit form of the relation in (1.46) for writing the path-ordered integral is called the Magnus formula (see Appendix A). If the commutators of $: \mathcal{H} :$ at different values of z, or equivalently the Poisson brackets of $\mathcal{H}$ at different values of z, are zero, i.e.,

$$\begin{aligned}
\left[: \mathcal{H}(z_1) : , : \mathcal{H}(z_2) :\right] &\equiv : \left\{\mathcal{H}(z_1) , \mathcal{H}(z_2)\right\} : \\
&= 0 , \qquad \forall z_1, z_2 \in \left(z_{\rm in}, z_{\rm out}\right) ,
\end{aligned} \tag{1.47}$$

then, the z-ordering in (1.46) is redundant and the transfer map reduces to the simple expression

$$\mathcal{M} = \exp\left(-\int_{z_{\rm in}}^{z_{\rm out}} dz : \mathcal{H}(z) :\right). \tag{1.48}$$

In many physical situations the z-ordering may be ignored and $\mathcal{M}$ can be approximated as in (1.48). It is clear that a knowledge of the transfer map is equivalent to the knowledge of the trajectories generated by the Hamiltonian $\mathcal{H}$. In the rest of the section we shall be illustrating certain basic results which facilitate the computation of the transfer maps, products of transfer maps and other related results, using the Lie algebraic structure underlying the Hamiltonian mechanics as exhibited above.

Before going into the details of the computation let us look at the case of the free particle. For a free particle $\phi = 0$ and $\boldsymbol{A} = (0, 0, 0)$. Consequently the Hamiltonian in (1.13) reduces to

$$\mathcal{H}_F = -\sqrt{p_0^2 - p_\perp^2} \tag{1.49}$$

and the transfer map in (1.46) becomes, for any Δz

$$\mathcal{M}_F = \exp\left(: \Delta z \sqrt{p_0^2 - p_\perp^2} :\right). \tag{1.50}$$

It is straightforward to get the results

$$\begin{aligned} \boldsymbol{r}_{\perp,{\rm out}} &= (\mathcal{M}_F \boldsymbol{r}_\perp)_{\rm in} = \boldsymbol{r}_{\perp,{\rm in}} + \Delta z \frac{\boldsymbol{p}_{\perp,{\rm in}}}{\sqrt{p_0^2 - p_{\perp,{\rm in}}^2}} \\ &\approx \boldsymbol{r}_{\perp,{\rm in}} + \Delta z \frac{\boldsymbol{p}_{\perp,{\rm in}}}{p_0} \qquad \text{(paraxial approximation)} \end{aligned} \tag{1.51}$$

and

$$\boldsymbol{p}_{\perp,{\rm out}} = (\mathcal{M}_F \boldsymbol{p}_\perp)_{\rm in} = \boldsymbol{p}_{\perp,{\rm in}}. \tag{1.52}$$

The above relations for the paraxial case can be compactly written in the matrix form as

$$\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm out} = M_F \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} = \begin{pmatrix} 1 & \Delta z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in}. \tag{1.53}$$

The above relation is immediately recognized as the relation for a drift through a distance Δz.

In the free particle case the computation can be done exactly. We shall resort to the example of free particle case at times for guidance to build the formalism and at other times to use it as the first example to demonstrate the formalism. In general, the transfer map is an infinite product of simpler maps and normally one uses a suitable truncated expression. In order to understand this we need the following theorem obtained in the context of Hamiltonian mechanics in which the transfer maps are symplectic preserving the Poisson brackets between $\boldsymbol{r}$ and $\boldsymbol{p}$.

Theorem: (Dragt-Finn) [28] Suppose that $\mathcal{M}$ is any symplectic map that sends the origin of the phase-space into itself, i.e., if $\boldsymbol{w}_{\rm in} = \mathbf{o}$ then $\boldsymbol{w}_{\rm out} = \mathbf{o}$. Then, $\mathcal{M}$ can be factored as a product of Lie transformations in the form

$$\begin{aligned} \mathcal{M} &= \mathcal{M}_2\mathcal{M}_3\mathcal{M}_4 \cdots \mathcal{M}_n \cdots \\ &= \exp\left(:\mathcal{F}_2:\right)\, \exp\left(:\mathcal{F}_3:\right)\, \exp\left(:\mathcal{F}_4:\right) \cdots \exp\left(:\mathcal{F}_n:\right) \cdots \end{aligned} \tag{1.54}$$

where each $\mathcal{F}_n$ is a homogeneous polynomial of degree n in the components of $\boldsymbol{w}$. Moreover the map is symplectic for any set of polynomials. Finally if the product is truncated at any stage the result is still a symplectic map.

In many a situation it is more convenient to use a factorization in an order opposite to that of the one given in the theorem above: i.e.,

$$\mathcal{M} = \cdots \exp\left(:\mathcal{G}_n:\right) \cdots \exp\left(:\mathcal{G}_4:\right)\, \exp\left(:\mathcal{G}_3:\right) \exp\left(:\mathcal{G}_2:\right). \tag{1.55}$$

It may be noted that in (1.55) the paraxial part of the map will not be affected by the reverse order of factorization, i.e., $\mathcal{G}_2 \equiv \mathcal{F}_2$; all the other $\mathcal{G}_n$ s are generally different from the corresponding $\mathcal{F}_n$ s. There are standard recipes available to switch from one order of factorization to the other.

Before proceeding further about the Lie algebraic transforms we consider a few examples arising out of the general paraxial Hamiltonian $\mathcal{H}_p$ given in (1.21).

Free Drift: This corresponds to the simplest possible Lie transformation

$$\begin{aligned}\mathcal{M}_D &= \exp\left(:\mathcal{F}_D:\right)\\ \mathcal{F}_D &= -\frac{\Delta z}{2p_0}p_\perp^2\,,\end{aligned} \tag{1.56}$$

obtained from the paraxial approximation of $\mathcal{M}_F$ in (1.50). Then,

$$\boldsymbol{r}_{\perp,\mathrm{out}} = \left(\mathcal{M}_D\boldsymbol{r}_\perp\right)_{\mathrm{in}} = \boldsymbol{r}_{\perp,\mathrm{in}} + \frac{\Delta z}{p_0}\boldsymbol{p}_{\perp,\mathrm{in}} \tag{1.57}$$

and

$$\boldsymbol{p}_{\perp,\mathrm{out}} = \left(\mathcal{M}_D\boldsymbol{p}_\perp\right)_{\mathrm{in}} = \boldsymbol{p}_{\perp,\mathrm{in}}\,. \tag{1.58}$$

The above relations can be compactly written in matrix form as

$$\begin{pmatrix}\boldsymbol{r}_\perp\\ \boldsymbol{p}_\perp/p_0\end{pmatrix}_{\mathrm{out}} = M_D\begin{pmatrix}\boldsymbol{r}_\perp\\ \boldsymbol{p}_\perp/p_0\end{pmatrix}_{\mathrm{in}} = \begin{pmatrix}1 & \Delta z\\ 0 & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{r}_\perp\\ \boldsymbol{p}_\perp/p_0\end{pmatrix}_{\mathrm{in}}. \tag{1.59}$$

The above relation is immediately recognized as the relation for a drift through a distance Δz.

Simple Thin Lens: For this system the Lie transformation is:

$$\begin{aligned}\mathcal{M}_L &= \exp\left(:\mathcal{F}_L:\right)\\ \mathcal{F}_L &= -\frac{p_0}{2f}r_\perp^2\,.\end{aligned} \tag{1.60}$$

This corresponds to the action of a thin lens with focal length f. Like in the previous example we look at the action of the transformation map:

$$\boldsymbol{r}_{\perp,\mathrm{out}} = \left(\mathcal{M}_L\boldsymbol{r}_\perp\right)_{\mathrm{in}} = \boldsymbol{r}_{\perp,\mathrm{in}} \tag{1.61}$$

and

$$\boldsymbol{p}_{\perp,\mathrm{out}} = \left(\mathcal{M}_L\boldsymbol{p}_\perp\right)_{\mathrm{in}} = -\frac{p_0}{f}\boldsymbol{r}_{\perp,\mathrm{in}} + \boldsymbol{p}_{\perp,\mathrm{in}}\,. \tag{1.62}$$

The above relations can be compactly written in matrix form as

$$\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm out} = M_L \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} . \tag{1.63}$$

The above relation is the familiar transfer matrix for a thin lens of focal length f.

Rotation in Magnetic Lenses: Image rotation in magnetic lenses correspond to the Lie transformation

$$\begin{aligned} \mathcal{M}_R &= \exp\left(: \mathcal{F}_R :\right) \\ \mathcal{F}_R &= -\theta L_z . \end{aligned} \tag{1.64}$$

The result of this transformation is easily seen to be

$$\begin{pmatrix} x \\ y \end{pmatrix}_{\rm out} = M_R \begin{pmatrix} x \\ y \end{pmatrix}_{\rm in} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}_{\rm in} \tag{1.65}$$

and

$$\begin{pmatrix} p_x \\ p_y \end{pmatrix}_{\rm out} = M_R \begin{pmatrix} p_x \\ p_y \end{pmatrix}_{\rm in} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} p_x \\ p_y \end{pmatrix}_{\rm in} \tag{1.66}$$

describing the familiar relations for rotation by an angle θ around the z-axis.

Imaging and Telescoping: Let us now consider the Lie transformation corresponding to the term $\boldsymbol{r}_\perp \cdot \boldsymbol{p}_\perp$ of $\mathcal{H}_p$ in (1.21). In this case

$$\begin{aligned} \mathcal{M}_{IT} &= \exp\left(: \mathcal{F}_{IT} :\right) \\ \mathcal{F}_{IT} &= \pm\sigma(\boldsymbol{r}_\perp \cdot \boldsymbol{p}_\perp) . \end{aligned} \tag{1.67}$$

Then, one has

$$\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm out} = M_{IT} \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} = \begin{pmatrix} e^{\mp\sigma} & 0 \\ 0 & e^{\pm\sigma} \end{pmatrix} \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} . \tag{1.68}$$

Such a system is both imaging and telescopic.

In the forgoing examples we have seen that the phase-space transfer maps corresponding to the key features (free drift, lens action, image rotation, and imaging/telescoping property) of any paraxial optical system can be individually expressed through the transfer maps ($\mathcal{M}_D$, $\mathcal{M}_L$, $\mathcal{M}_R$ and $\mathcal{M}_{IT}$) represented as Lie transformations. Thus, one describes paraxial optics in terms of maps (say $\mathcal{M}_p$ or $\mathcal{M}_2$) associated with the quadratic function (say $\mathcal{F}_p$ or $\mathcal{F}_2$). From the traditional matrix methods of paraxial optics we know that the net effect of combinations of individual optical elements can be expressed by taking the product of the corresponding transfer matrices. In the operator approach in terms of Lie transformations the net effect of any optical system consisting of several parts is obtained by the product of the corresponding transformation operators.

To illustrate the map corresponding to a complete paraxial Hamiltonian let us consider the axially symmetric magnetic lens. In this case, the paraxial Hamiltonian is

$$\begin{aligned} \mathcal{H}_p &= -p_0 + \mathcal{H}_0 + \mathcal{H}_R \\ \mathcal{H}_0 &= \frac{1}{2p_0} p_\perp^2 + \frac{p_0}{2} \alpha^2 r_\perp^2 \qquad (1.69) \\ \mathcal{H}_R &= -\alpha L_z\,. \qquad (1.70) \end{aligned}$$

as seen already ((1.26)-(1.29)). Noting that

$$\{\mathcal{H}_0\,, \mathcal{H}_R\} = 0\,, \qquad (1.71)$$

we have

$$\mathcal{M}_p = \mathcal{M}_R \mathcal{M}_0\,, \qquad (1.72)$$

where

$$\mathcal{M}_R = \exp\left(\int_{z_{in}}^{z} dz : \alpha(z) L_z :\right) = \exp\left(: \theta(z, z_{\rm in}) L_z :\right) \qquad (1.73)$$

with $\theta(z\,, z_{\rm in}) = \int_{z_{\rm in}}^{z} dz \alpha(z)$ as the angle by which the image is rotated around the z-axis.

From the earlier discussion we know that $\mathcal{M}_0$ is given by

$$\mathcal{M}_0 = \wp \left\{ \exp\left(- \int_{z_{\rm in}}^{z_{\rm out}} dz : \frac{1}{2p_0} p_\perp^2 + \frac{p_0}{2} \alpha(z)^2 r_\perp^2 : \right) \right\}. \qquad (1.74)$$

One can show that the effect of this Lie transformation can be written as

$$
\begin{aligned}
\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\text{out}} &= M_0 \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\text{in}} \\
&= \begin{pmatrix} g(z) & h(z) \\ g'(z) & h'(z) \end{pmatrix} \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\text{in}}
\end{aligned}
\tag{1.75}
$$

where $g(z)$ and $h(z)$ are the two linearly independent solutions of the classical trajectory equations subject to the initial conditions

$$
\begin{aligned}
g(z_{\text{in}}) &= h'(z_{\text{in}}) = 1 \\
g'(z_{\text{in}}) &= h(z_{\text{in}}) = 0\,,
\end{aligned}
\tag{1.76}
$$

and the symplecticity condition

$$
g(z)h'(z) - h(z)g'(z) = 1\,, \qquad \forall\ z\,. \tag{1.77}
$$

The arguments leading to the above result are parallel to those to be presented in the quantum case (with commutator brackets replacing the Poisson brackets) later where a series method for getting the solutions $g(z)$ and $h(z)$ will also emerge. For the present it suffices to note that the above result is the basis for understanding the Gaussian imaging by a magnetic lens from the point of view of classical mechanics, or geometrical charged-particle optics.

So far we have examined only the paraxial case. Any meaningful system has departures from the paraxial approximation and one has to deal with the Hamiltonian with more than quadratic terms in $(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp)$. To proceed further, we have to make use of the interaction picture description introduced by Dragt *et al.* in this context adapting the well-known interaction picture of quantum dynamics to deal with time-dependent perturbations. The basic relations required to obtain the interaction picture are the factorization theorem, Magnus formula and BCH-formula.

We just state them here. If $\hat{A}$ and $\hat{B}$ are two noncommuting operators, then,

$$e^{\hat{A}+\hat{B}} = \cdots e^{\hat{C}_n} \cdots e^{\left(\frac{1}{6}[\hat{A},[\hat{A},\hat{B}]]+\frac{1}{3}[\hat{B},[\hat{A},\hat{B}]]\right)} e^{\left(\frac{1}{2}[\hat{A},\hat{B}]\right)} e^{\hat{B}} e^{\hat{A}} \tag{1.78}$$

where we we choose $\deg(\hat{B}) \geq \deg(\hat{A})$ and $\hat{C}_n$ is a commutator of rank n. If the Hamiltonian of a system is given by

$$\mathcal{H} = \mathcal{H}_p + \mathcal{H}_a\,, \tag{1.79}$$

then, the corresponding transfer operator $\mathcal{M}$ has the form

$$\mathcal{M} = \mathcal{M}_p \mathcal{M}_a = \exp\left(:\mathcal{F}_p:\right)\exp\left(:\mathcal{F}_a:\right)\,. \tag{1.80}$$

Now, we can rewrite this equation as

$$\begin{aligned} \mathcal{M} &= \left(\mathcal{M}_p \mathcal{M}_a \mathcal{M}_p^{-1}\right)\mathcal{M}_p \\ &= \exp\left(:\exp\left(:\mathcal{F}_p:\right)\mathcal{F}_a:\right)\exp\left(:\mathcal{F}_p:\right) \\ &= \exp\left(:\mathcal{F}_a^I:\right)\exp\left(:\mathcal{F}_p:\right) = \mathcal{M}_a^I \mathcal{M}_p\,, \end{aligned} \tag{1.81}$$

where the superscript I stands for interaction picture map. Explicitly writing

$$\exp\left(:\mathcal{F}_a^I:\right) = \wp\left\{\exp\left(-\int_{z_{in}}^{z_{out}} dz\ :\ \mathcal{H}_a^I(\boldsymbol{w}, z)\ :\right)\right\}, \tag{1.82}$$

with

$$\mathcal{H}_a^I(\boldsymbol{w}, z) = \mathcal{H}_a(\mathcal{M}_p \boldsymbol{w}, z)\,. \tag{1.83}$$

Then, using the factorization theorem mentioned above one can write

$$\begin{aligned} \mathcal{M} &= \mathcal{M}_a^I \mathcal{M}_p \\ \mathcal{M}_a^I &= \cdots \exp\left(:\mathcal{G}_n^I:\right) \cdots \exp\left(:\mathcal{G}_4^I:\right)\ \exp\left(:\mathcal{G}_3^I:\right), \end{aligned} \tag{1.84}$$

where

$$\mathcal{G}_3^I = -\int_{z_{in}}^{z_{out}} dz \mathcal{H}_{(3)}^I(z)$$

$$
\begin{aligned}
\mathcal{G}_4^I &= -\int_{z_{\rm in}}^{z_{\rm out}} dz \mathcal{H}_{(4)}^I(z) - \frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz_1 \int_{z_{\rm in}}^{z_1} dz_2 \left\{\mathcal{H}_{(3)}^I(z_1), \mathcal{H}_{(3)}^I(z_2)\right\} \\
\mathcal{G}_5^I &= -\int_{z_{\rm in}}^{z_{\rm out}} dz \mathcal{H}_{(5)}^I(z) + \int_{z_{\rm in}}^{z_{\rm out}} dz_1 \int_{z_{\rm in}}^{z_1} dz_2 \left\{\mathcal{H}_{(3)}^I(z_1), \mathcal{H}_{(4)}^I(z_2)\right\} \\
&\qquad -\frac{1}{3}\int_{z_{\rm in}}^{z_{\rm out}} dz_1 \int_{z_{\rm in}}^{z_1} dz_2 \int_{z_{\rm in}}^{z_2} dz_3 \left(\left\{\mathcal{H}_{(3)}^I(z_3), \left\{\mathcal{H}_{(3)}^I(z_2), \mathcal{H}_{(3)}^I(z_1)\right\}\right\}\right. \\
&\qquad\qquad \left. + \left\{\mathcal{H}_{(3)}^I(z_2), \left\{\mathcal{H}_{(3)}^I(z_3), \mathcal{H}_{(3)}^I(z_1)\right\}\right\}\right) \\
&\vdots \\
\mathcal{G}_n^I &= -\int_{z_{\rm in}}^{z_{\rm out}} dz \mathcal{H}_{(n)}^I(z) \qquad + \text{ multiple Poisson bracket terms}. \qquad (1.85)
\end{aligned}
$$

Expressions for $\mathcal{G}_n^I$ can be obtained up to any desired order of accuracy. For the present purpose we need only up to $n = 4$.

In the case of the magnetic round lens we have

$$
\begin{aligned}
\mathcal{G}_4^I = -\Bigg(& \frac{1}{4p_0^3} C p_\perp^4 \\
& + \frac{1}{2p_0^2} K p_\perp^2 \left(p_\perp \cdot r_\perp + r_\perp \cdot p_\perp\right) \\
& + \frac{1}{p_0^2} k p_\perp^2 L_z \\
& + \frac{1}{p_0} A \left(r_\perp \cdot p_\perp\right)^2 \\
& + \frac{1}{p_0} a \left(r_\perp \cdot p_\perp\right) \hat{L}_z \\
& + \frac{1}{2p_0} F r_\perp^2 p_\perp^2 \\
& + D r_\perp^2 \left(r_\perp^2 \cdot p_\perp^2\right) \\
& + d r_\perp^2 L_z \\
& + \frac{p_0}{4} E r_\perp^4 \Bigg), \qquad (1.86)
\end{aligned}
$$

where

$$
\begin{aligned}
C &= \frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz \left\{\left(\alpha^4 - \alpha\alpha''\right) h^4 + 2\alpha^2 h^2 h'^2 + h'^4\right\} \\
K &= \frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz \left\{\left(\alpha^4 - \alpha\alpha''\right) g h^3 + \alpha^2 (gh)' h h' + g' h'^3\right\}
\end{aligned}
$$

$$
\begin{aligned}
k &= \int_{z_{\rm in}}^{z_{\rm out}} dz \left\{ \left(\frac{1}{8}\alpha'' - \frac{1}{2}\alpha^3 \right) h^2 - \frac{1}{2}\alpha h'^2 \right\} \\
A &= \frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz \left\{ \left(\alpha^4 - \alpha\alpha'' \right) g^2 h^2 + 2\alpha^2 g g' h h' + g'^2 h'^2 - \alpha^2 \right\} \\
a &= \int_{z_{\rm in}}^{z_{\rm out}} dz \left\{ \left(\frac{1}{4}\alpha'' - \alpha^3 \right) gh - \alpha g' h' \right\} \\
F &= \frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz \left\{ \left(\alpha^4 - \alpha\alpha'' \right) g^2 h^2 + \alpha^2 \left(g^2 h'^2 + g'^2 h^2 \right) + g'^2 h'^2 + 2\alpha^2 \right\} \\
D &= \frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz \left\{ \left(\alpha^4 - \alpha\alpha'' \right) g^3 h + \alpha^2 g g' (gh)' + g'^3 h' \right\} \\
d &= \int_{z_{\rm in}}^{z_{\rm out}} dz \left\{ \left(\frac{1}{8}\alpha'' - \frac{1}{2}\alpha^3 \right) g^2 - \frac{1}{2}\alpha g'^2 \right\} \\
E &= \frac{1}{2}\int_{z_{\rm in}}^{z_{\rm out}} dz \left\{ \left(\alpha^4 - \alpha\alpha'' \right) g^4 + 2\alpha g^2 g'^2 + g'^4 \right\} . \qquad (1.87)
\end{aligned}
$$

The constants C, K, k, A, a, F, D and d are the well-known aberration coefficients corresponding to spherical aberration, coma, anisotropic coma, astigmatism, anisotropic astigmatism, curvature of field, distortion and anisotropic distortion, respectively, when $z_{\rm in}$ and $z_{\rm out}$ represent the coordinates of the object plane and the image plane respectively. We shall discuss the aberrations in detail in Chapter II based on quantum mechanics.

Chapter 2

Scalar quantum theory of charged-particle beam optics

2.1 Formalism

This chapter is devoted to the development of a general formalism of scalar quantum theory of charged-particle beam optics, for situations when the spin is zero or can be assumed to be just a spectator. To illustrate the general formalism we consider its applications to the examples of free propagation (diffraction), round magnetic lens (electron-optical imaging), magnetic quadrupole lens (accelerator optics) and electrostatic round and quadrupole lenses. The starting point for such a formalism is of course the basic equations of scalar quantum mechanics namely, the Schrödinger and Klein-Gordon equations for the nonrelativistic and relativistic cases respectively. As in the case of the classical theory of charged-particle beam transport, where the basic Hamiltonian of classical mechanics was cast into a beam-optical form, here too the first step lies in casting the above basic equations into beam-optical forms. It will be seen shortly that the beam-optical form of both Schrödinger and Klein-Gordon equations are identical except for the interpretation of the design momentum as nonrelativistic or relativistic respectively.

The nonrelativistic Schrödinger equation for a particle of charge q and mass m_0

moving in a static electromagnetic field with potentials $(\phi(\boldsymbol{r}), \boldsymbol{A}(\boldsymbol{r}))$ is

$$\left(\mathrm{i}\hbar\frac{\partial}{\partial t} - q\phi\right)\Psi(\boldsymbol{r},t) = \frac{1}{2m_0}\hat{\pi}^2\Psi(\boldsymbol{r},t)\,, \tag{2.1}$$

with the usual notations

$$\hat{\boldsymbol{\pi}} = \hat{\boldsymbol{p}} - q\boldsymbol{A}(\boldsymbol{r})\,, \quad \hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla\,, \quad \hat{\pi}^2 = \hat{\pi}_x^2 + \hat{\pi}_y^2 + \hat{\pi}_z^2\,. \tag{2.2}$$

Since we are dealing with time-independent systems the wavefunctions of the particles constituting a monoenergetic beam take the form

$$\Psi(\boldsymbol{r},t) = \mathrm{e}^{-\mathrm{i}Et/\hbar}\psi(\boldsymbol{r})\,. \tag{2.3}$$

Then equation (2.1) becomes

$$\left(\hat{\pi}^2 - |\boldsymbol{\pi}|^2\right)\psi(\boldsymbol{r}) = 0\,, \tag{2.4}$$

where $|\boldsymbol{\pi}| = \sqrt{2m_0\,(E - q\phi)}$ is the nonrelativistic value for the kinetic momentum. Now, introducing the notation

$$\hat{\boldsymbol{D}} = \frac{\mathrm{i}}{\hbar}\hat{\boldsymbol{\pi}} = \left(\nabla - \frac{\mathrm{i}}{\hbar}q\boldsymbol{A}\right)\,, \tag{2.5}$$

and denoting by $k(\boldsymbol{r})$ the wavenumber of the particle in accordance with the non-relativistic de Broglie relation,

$$k(\boldsymbol{r}) = |\boldsymbol{\pi}|/\hbar = \sqrt{2m_0\,(E - q\phi(\boldsymbol{r}))}/\hbar\,, \tag{2.6}$$

the stationary form of the Schrödinger equation (2.4) takes the Helmholtz-like form

$$\{\hat{D}^2 + k^2(\boldsymbol{r})\}\psi(\boldsymbol{r}) = 0\,. \tag{2.7}$$

TheKlein-Gordon equation

$$\left(\mathrm{i}\hbar\frac{\partial}{\partial t} - q\phi\right)^2\Psi(\boldsymbol{r},t) = \{c^2\hat{\pi}^2 + m_0^2c^4\}\Psi(\boldsymbol{r},t) \tag{2.8}$$

can also be cast in the same Helmholtz-like form (2.7) with the only significant difference that the expression for $k(\boldsymbol{r})$ will be the relativistic one. Now, let us take the form of the wavefunction of the particle of the monoenergetic beam to be of the form

$$\Psi(\boldsymbol{r}, t) = \mathrm{e}^{-\mathrm{i}(m_0c^2+E)t/\hbar}\psi(\boldsymbol{r})\,. \tag{2.9}$$

where E is the dynamic (kinetic + potential) part of the total energy of the particle with m_0c^2 being the rest energy. Then, the Klein-Gordon equation (2.8) takes the stationary form

$$\left\{\hat{\pi}^2 - 2Em_0\left(1 + \frac{E}{2m_0c^2}\right) + 2qm_0\phi\left(1 + \frac{E}{m_0c^2} - \frac{q\phi}{2m_0c^2}\right)\right\}\psi(\boldsymbol{r}) = 0 \tag{2.10}$$

which can be equivalently written as

$$\{\hat{D}^2 + k^2(\boldsymbol{r})\}\psi(\boldsymbol{r}) = 0\,, \tag{2.11}$$

where

$$k^2(\boldsymbol{r}) = \frac{1}{\hbar^2}\left\{2Em_0\left(1 + \frac{E}{2m_0c^2}\right) - 2qm_0\phi(\boldsymbol{r})\left(1 + \frac{E}{m_0c^2} - \frac{q\phi(\boldsymbol{r})}{2m_0c^2}\right)\right\}\,, \tag{2.12}$$

consistent with the relativistic Einstein-de Broglie relation

$$E = \{c^2\hbar^2k^2(\boldsymbol{r}) + m_0^2c^4\}^{\frac{1}{2}} - m_0c^2 + q\phi(\boldsymbol{r})\,. \tag{2.13}$$

Thus it is shown that equation (2.11) based on the relativistic Klein-Gordon equation has the same Helmholtz-like form as equation (2.7) based on the nonrelativistic Schrödinger equation; the only difference is in the expressions for $k(\boldsymbol{r})$ in the two cases.

Equation (2.11) is the basic equation to be used as the starting point to develop a complete scalar quantum theory of charged-particle beam transport through any optical system. It is to be emphasized that so far we have essentially cast the time-dependent Schrödinger and Klein-Gordon equations for beams propagating with a single conserved energy (elastic scattering by stationary electromagnetic field constituting the optical system) into a Helmholtz-like form by eliminating the time-dependence, without making any assumption on the geometry of the system to be studied. To proceed further, the system has to be chosen a bit more specifically. So, we shall consider the system to have a straight optic axis along the z-axis of a Cartesian coordinate frame and consider the monoenergetic charged-particle beam to be quasiparaxial and moving close to the $+z$-direction. Let it be assumed that the system is located between the xy-planes with the z-coordinates $z_{\rm in}$ and $z_{\rm out}$, i.e., the system field is practically zero in the 'input' region ($z \leq z_{\rm in}$) and the 'output' region ($z \geq z_{\rm out}$). By 'input' and 'output' regions we mean the regions outside the system and close to it. The constant wavenumber of the incident particle in the input region is given by

$$
\begin{aligned}
k(\boldsymbol{r}_\perp, z < z_{\rm in}) &= k_0 = \frac{p_0}{\hbar} \\
&= \frac{1}{\hbar c}\sqrt{E(E + 2m_0c^2)} = \frac{1}{\hbar}\left\{2Em_0\left(1 + \frac{E}{2m_0c^2}\right)\right\}^{\frac{1}{2}} , \qquad (2.14)
\end{aligned}
$$

as is seen by putting $\phi(\boldsymbol{r}) = 0$ in (2.12). After elastic scattering by the system the particle will emerge in the field-free output region with the same value of the wavenumber, namely, k_0 i.e., $k(\boldsymbol{r}_\perp, z > z_{\rm out})$ also has the value k_0. Since the beam is supposed to be monoenergetic, the wavevector $\boldsymbol{k}_0$ of any particle of it will have the same magnitude

$$
|\boldsymbol{k}_0| = \left(k_{0x}^2 + k_{0y}^2 + k_{0z}^2\right)^{\frac{1}{2}} = k_0 . \qquad (2.15)
$$

irrespective of its direction. Quasiparaxiality of the input beam implies that

$$
k_{0x}^2 + k_{0y}^2 = k_{0\perp}^2 \ll k_{0z}^2 . \qquad (2.16)
$$

Since we consider the optical system to be such that the input 'beam' emerges in the output region again as a 'beam', the relation

$$k_x(\boldsymbol{r})^2 + k_y(\boldsymbol{r})^2 = k_\perp^2(\boldsymbol{r}) \ll k_z(\boldsymbol{r})^2 \tag{2.17}$$

will be assumed to hold throughout the propagation of the beam. Further, since we are always concerned only with the forward propagating beam close to the $+z$-direction the beam wavefunction we consider, throughout the transport of the beam, would be a packet, or linear combination, of only those plane waves corresponding to wavevectors satisfying the conditions

$$k_\perp^2 \ll k_z^2 , \qquad k_z > 0 . \tag{2.18}$$

Our aim is to relate the beam wavefunction in the field-free output region,

$$\psi_{\mathrm{out}}(\boldsymbol{r}) = \psi(\boldsymbol{r}_\perp , z \geq z_{\mathrm{out}}) , \tag{2.19}$$

to the beam wavefunction in the field-free input region,

$$\psi_{\mathrm{in}}(\boldsymbol{r}) = \psi(\boldsymbol{r}_\perp , z \leq z_{\mathrm{in}}) , \tag{2.20}$$

so that the values of the observable beam characteristics in the output region can be related to their values in the input region using the wavefunction. To this end, the most desirable starting point would be an z-evolution equation for $\psi(\boldsymbol{r}_\perp , z)$ linear in $\frac{\partial}{\partial z}$. So, first, we cast equation (2.12) in such a form using a method similar to the way in which the Klein-Gordon equation is written in the Feshbach-Villars form (linear in $\frac{\partial}{\partial t}$) [29], unlike the Klein-Gordon equation (quadratic in $\frac{\partial}{\partial t}$). (see Appendix B for the Feshbach-Villars form of the Klein-Gordon equation.)

Let us examine the expression for $k^2(\boldsymbol{r})$ in (2.12), in some detail. It can be partitioned in the following way,

$$\begin{aligned}
k^2(\boldsymbol{r}) &= k_0^2 - \tilde{k}^2(\boldsymbol{r}) \\
k_0^2 &= \frac{1}{\hbar^2}\left\{2Em_0\left(1 + \frac{E}{2m_0c^2}\right)\right\} \\
\tilde{k}^2(\boldsymbol{r}) &= \frac{1}{\hbar^2}\left\{2qm_0\phi(\boldsymbol{r})\left(1 + \frac{E}{m_0c^2} - \frac{q\phi(\boldsymbol{r})}{2m_0c^2}\right)\right\} .
\end{aligned} \tag{2.21}$$

The first point to be noted is that k_0^2 is constant, independent of the electric scalar potential $\phi(\boldsymbol{r})$, and depends only on the kinetic energy E of the incident beam-particle, where as $\tilde{k}^2(\boldsymbol{r})$ depends on both the incident kinetic energy E and the potential $\phi(\boldsymbol{r})$. For magnetic lenses $\phi(\boldsymbol{r})$ is zero and consequently $k^2(\boldsymbol{r}) \equiv k_0^2$.

Now equation (2.11) takes the form

$$\{\hat{D}^2 + k_0^2 - \tilde{k}^2(\boldsymbol{r})\}\psi(\boldsymbol{r}) = 0\,. \tag{2.22}$$

As stated earlier the above equation is the starting point for the scalar quantum theory of charged-particle beam optics. The first thing to be noted is that the above Helmholtz-like form is quadratic in $\frac{\partial}{\partial z}$ where as the most desired form would be an z-evolution equation for the wavefunction linear in $\frac{\partial}{\partial z}$. This is achieved by adopting a method similar to the method in which the scalar Klein-Gordon equation (quadratic in $\frac{\partial}{\partial t}$) is written in the two-component Feshbach-Villars form (linear in $\frac{\partial}{\partial t}$). This provides a Feshbach-Villars-like two-component form for the standard Klein-Gordon equation, now with z, the coordinate along the optic axis playing the role of t. (see Appendix B for the Feshbach-Villars form of the Klein-Gordon equation.)

To this end, let

$$\left(\begin{array}{c} \psi_1(\boldsymbol{r}) \\ \psi_2(\boldsymbol{r}) \end{array}\right) = \left(\begin{array}{c} \psi(\boldsymbol{r}) \\ -\frac{\mathrm{i}}{k_0}\left(\frac{\partial}{\partial z} - \frac{\mathrm{i}}{\hbar}qA_z(\boldsymbol{r})\right)\psi(\boldsymbol{r}) \end{array}\right). \tag{2.23}$$

Then, equation (2.22) is equivalent to

$$\begin{aligned} & -\frac{\mathrm{i}}{k_0}\left(\frac{\partial}{\partial z} - \frac{\mathrm{i}}{\hbar}qA_z\right)\left(\begin{array}{c} \psi_1 \\ \psi_2 \end{array}\right) \\ & \quad = \left(\begin{array}{cc} 0 & 1 \\ \frac{1}{k_0^2}\{k_0^2 - \tilde{k}^2 + \hat{D}_\perp^2\} & 0 \end{array}\right)\left(\begin{array}{c} \psi_1 \\ \psi_2 \end{array}\right). \end{aligned} \tag{2.24}$$

Next we make the transformation

$$\begin{aligned} \left(\begin{array}{c} \psi_1 \\ \psi_2 \end{array}\right) \longrightarrow \left(\begin{array}{c} \psi_+ \\ \psi_- \end{array}\right) &= M\left(\begin{array}{c} \psi_1 \\ \psi_2 \end{array}\right) = \frac{1}{2}\left(\begin{array}{c} \psi_1 + \psi_2 \\ \psi_1 - \psi_2 \end{array}\right) \\ M &= \frac{1}{2}\left(\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right) \qquad M^{-1} = \left(\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right) = 2M\,. \end{aligned} \tag{2.25}$$

Consequently, equation (2.24) can be written as

$$
\begin{aligned}
& -\frac{\mathrm{i}}{k_0}\left(\frac{\partial}{\partial z}-\frac{\mathrm{i}}{\hbar}qA_z\right)\left(\begin{array}{c}\psi_+\\ \psi_-\end{array}\right)\\
& \quad = -\frac{\mathrm{i}}{k_0}\left(\frac{\partial}{\partial z}-\frac{\mathrm{i}}{\hbar}qA_z\right)M\left(\begin{array}{c}\psi_1\\ \psi_2\end{array}\right)\\
& \quad = M\left\{-\frac{\mathrm{i}}{k_0}\left(\frac{\partial}{\partial z}-\frac{\mathrm{i}}{\hbar}qA_z\right)\right\}\left(\begin{array}{c}\psi_1\\ \psi_2\end{array}\right)\\
& \quad = M\left(\begin{array}{cc}0 & 1\\ \frac{1}{k_0^2}\{k_0^2-\tilde{k}^2+\hat{D}_\perp^2\} & 0\end{array}\right)\left(\begin{array}{c}\psi_1\\ \psi_2\end{array}\right)\\
& \quad = M\left(\begin{array}{cc}0 & 1\\ \frac{1}{k_0^2}\{k_0^2-\tilde{k}^2+\hat{D}_\perp^2\} & 0\end{array}\right)M^{-1}\left(\begin{array}{c}\psi_+\\ \psi_-\end{array}\right)\\
& \quad = \left(\begin{array}{cc}1+\frac{1}{2k_0^2}(\hat{D}_\perp^2-\tilde{k}^2) & \frac{1}{2k_0^2}(\hat{D}_\perp^2-\tilde{k}^2)\\ -\frac{1}{2k_0^2}(\hat{D}_\perp^2-\tilde{k}^2) & -1-\frac{1}{2k_0^2}(\hat{D}_\perp^2-\tilde{k}^2)\end{array}\right)\left(\begin{array}{c}\psi_+\\ \psi_-\end{array}\right).
\end{aligned}
\tag{2.26}
$$

Rearranging the equation (2.26) we get

$$
\frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\left(\begin{array}{c}\psi_+\\ \psi_-\end{array}\right)=\hat{\mathrm{H}}\left(\begin{array}{c}\psi_+\\ \psi_-\end{array}\right),
\tag{2.27}
$$

with

$$
\begin{aligned}
\hat{\mathrm{H}} &= -\sigma_z+\hat{\mathcal{E}}+\hat{\mathcal{O}} && (2.28)\\
\hat{\mathcal{E}} &= -\frac{q}{\hbar k_0}A_z\mathbb{1}-\frac{1}{2k_0^2}(\hat{D}_\perp^2-\tilde{k}^2)\sigma_z && (2.29)\\
\hat{\mathcal{O}} &= -\frac{\mathrm{i}}{2k_0^2}(\hat{D}_\perp^2-\tilde{k}^2)\sigma_y\,, && (2.30)
\end{aligned}
$$

where $\mathbb{1}$ is the 2×2 identity matrix and σ_y and σ_z are, respectively, the y and z components of the triplet of Pauli matrices

$$
\sigma=\left(\sigma_x=\left(\begin{array}{cc}0 & 1\\ 1 & 0\end{array}\right),\ \sigma_y=\left(\begin{array}{cc}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{array}\right),\ \sigma_z=\left(\begin{array}{cc}1 & 0\\ 0 & -1\end{array}\right)\right).
\tag{2.31}
$$

In order to understand equation (2.27) better, let us see what it means in the case of propagation of the beam through free space. In free space, with $\phi=0$ and $\boldsymbol{A}=(0,0,0)$, we have $\tilde{k}^2(\boldsymbol{r})=0$ and

$$
\frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\left(\begin{array}{c}\psi_+\\ \psi_-\end{array}\right)=\left(\begin{array}{cc}-1-\frac{1}{2k_0^2}\nabla_\perp^2 & -\frac{1}{2k_0^2}\nabla_\perp^2\\ \frac{1}{2k_0^2}\nabla_\perp^2 & 1+\frac{1}{2k_0^2}\nabla_\perp^2\end{array}\right)\left(\begin{array}{c}\psi_+\\ \psi_-\end{array}\right),
\tag{2.32}
$$

where $\nabla_\perp^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. A plane wave with a given $\boldsymbol{k}_0 = (k_{0x}, k_{0y}, k_{0z})$, namely,

$$\psi_{\boldsymbol{k}_0}(\boldsymbol{r}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}_0 \cdot \boldsymbol{r}}, \tag{2.33}$$

is associated with

$$\begin{aligned} \begin{pmatrix} \psi_{\boldsymbol{k}_0,+}(\boldsymbol{r}) \\ \psi_{\boldsymbol{k}_0,-}(\boldsymbol{r}) \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} \psi_{\boldsymbol{k}_0} - \frac{\mathrm{i}}{k_0} \frac{\partial \psi_{\boldsymbol{k}_0}}{\partial z} \\ \psi_{\boldsymbol{k}_0} + \frac{\mathrm{i}}{k_0} \frac{\partial \psi_{\boldsymbol{k}_0}}{\partial z} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 + \frac{k_{0z}}{k_0} \\ 1 - \frac{k_{0z}}{k_0} \end{pmatrix} \psi_{\boldsymbol{k}_0}(\boldsymbol{r}). \end{aligned} \tag{2.34}$$

It can be easily checked that this $\begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}$ satisfies equation (2.32). For a quasi-paraxial beam moving close to the $+z$-direction, with $k_{0z} > 0$ and $k_{0z} \approx k_0$, it is clear from (2.34) that $\psi_+ \gg \psi_-$. By extending this observation it can be seen easily that for any wavepacket of the form

$$\begin{aligned} \psi(\boldsymbol{r}) &= \int d^3k_0\, \varphi(\boldsymbol{k}_0) \psi_{\boldsymbol{k}_0}(\boldsymbol{r}), \qquad \int d^3k_0\, |\varphi(\boldsymbol{k}_0)|^2 = 1 \\ &\quad \text{with} \quad |\boldsymbol{k}_0| = k_0\,,\ k_{0z} \approx k_0\,,\ k_{0z} > 0\,, \end{aligned} \tag{2.35}$$

representing a monochromatic quasiparaxial beam moving close to the $+z$-direction,

$$\begin{aligned} \begin{pmatrix} \psi_+(\boldsymbol{r}) \\ \psi_-(\boldsymbol{r}) \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} \psi - \frac{\mathrm{i}}{k_0} \frac{\partial \psi}{\partial z} \\ \psi + \frac{\mathrm{i}}{k_0} \frac{\partial \psi}{\partial z} \end{pmatrix} \\ &= \begin{pmatrix} \int d^3k_0\, \varphi(\boldsymbol{k}_0) \psi_{\boldsymbol{k}_0,+}(\boldsymbol{r}) \\ \int d^3k_0\, \varphi(\boldsymbol{k}_0) \psi_{\boldsymbol{k}_0,-}(\boldsymbol{r}) \end{pmatrix}, \\ &\quad \text{with} \quad |\boldsymbol{k}_0| = k_0\,,\ k_{0z} \approx k_0\,,\ k_{0z} > 0\,, \end{aligned} \tag{2.36}$$

is such that $\psi_+(\boldsymbol{r}) \gg \psi_-(\boldsymbol{r})$. Thus, in general, in the representation of equation (2.27) we should expect ψ_+ to be large compared to ψ_- for any monochromatic quasiparaxial beam passing close to the $+z$-direction through the system supporting beam propagation.

To summarize, we have transformed the time-dependent Klein-Gordon equation into a time-independent form linear in $\frac{\partial}{\partial z}$. Further $\hat{\mathrm{H}}$ has been partitioned (apart from the leading term $-\sigma_z$), into an 'even part' $\hat{\mathcal{E}}$ and an 'odd part' $\hat{\mathcal{O}}$. The even part does not couple ψ_+ and ψ_- and the odd part couples them. We also make note of the algebraic property

$$\left[\sigma_z, \hat{\mathcal{E}}\right] = 0\,, \qquad \text{and} \qquad \left[\sigma_z, \hat{\mathcal{O}}\right] = 2\sigma_z\hat{\mathcal{O}}\,. \tag{2.37}$$

The motivation for following such a procedure comes from the Foldy-Wouthuysen technique originally developed for understanding the nonrelativistic limit of the Dirac equation [30] (see also [32]-[34]; see [31] where the technique has been applied to the Feshbach-Villars form of the Klein-Gordon equation). It will be shown below how the Foldy-Wouthuysen technique can be adopted for the quantum theory of charged-particle beam optics, both in the scalar (in this chapter) and the spinor case (to be covered in Chapter III) respectively in order to understand the behaviour of the optical elements in a systematic way starting with the paraxial approximation and considering the aberrations of the various orders one after the other. With this motivation let us first look at the Dirac equation

$$\frac{i\hbar}{m_0c^2}\frac{\partial}{\partial t}\left(\begin{array}{c}\Psi_u(\boldsymbol{r},t)\\ \Psi_l(\boldsymbol{r},t)\end{array}\right) = \hat{\mathrm{H}}_D\left(\begin{array}{c}\Psi_u(\boldsymbol{r},t)\\ \Psi_l(\boldsymbol{r},t)\end{array}\right)\,, \tag{2.38}$$

where

$$\Psi_u = \left(\begin{array}{c}\Psi_1\\ \Psi_2\end{array}\right)\,, \qquad \Psi_l = \left(\begin{array}{c}\Psi_3\\ \Psi_4\end{array}\right) \tag{2.39}$$

$$\hat{\mathrm{H}}_D = \beta + \hat{\mathcal{E}}_D + \hat{\mathcal{O}}_D \tag{2.40}$$

$$\hat{\mathcal{E}}_D = \frac{q\phi(\boldsymbol{r})}{m_0c^2} \tag{2.41}$$

$$\hat{\mathcal{O}}_D = \frac{1}{m_0c}\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}}\,, \tag{2.42}$$

$\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$ and β are the 4×4 Dirac matrices given by

$$\boldsymbol{\alpha} = \left(\begin{array}{cc}\mathrm{O} & \boldsymbol{\sigma}\\ \boldsymbol{\sigma} & \mathrm{O}\end{array}\right)\,, \quad \beta = \left(\begin{array}{cc}\mathbb{1} & \mathrm{O}\\ \mathrm{O} & -\mathbb{1}\end{array}\right)\,, \quad \mathrm{O} = \left(\begin{array}{cc}0 & 0\\ 0 & 0\end{array}\right)\,, \tag{2.43}$$

and Ψ_u and Ψ_l are the upper and lower pairs of components of the Dirac spinor wavefunction Ψ. The Dirac Hamiltonian $\hat{\mathrm{H}}_D$ has been partitioned (apart from the leading term β) into the 'even part' $\hat{\mathcal{E}}_D$ and an 'odd part' $\hat{\mathcal{O}}_D$ with the algebraic property

$$\left[\beta, \hat{\mathcal{E}}_D\right] = 0\,, \qquad \text{and} \qquad \left[\beta, \hat{\mathcal{O}}_D\right] = 2\beta\hat{\mathcal{O}}_D\,. \tag{2.44}$$

It is also well-known that for any positive energy Dirac Ψ in the nonrelativistic situation ($|\pi| \ll m_0c$) the upper components (Ψ_u) are large compared to the lower components (Ψ_l). The Foldy-Wouthuysen technique expands the Dirac Hamiltonian into a power series with $\frac{1}{m_0c}$ as the expansion parameter. Each successive approximation reduces the strength of the 'odd operator' by a factor of $\frac{1}{m_0c}$. In principle, one can reduce the strength of the odd operator to any desired degree of accuracy, of course, with the computations becoming more involved for greater degree of accuracy. Thus one obtains a systematic expansion of the Dirac equation leading to its nonrelativistic approximation plus a sequence of relativistic corrections (see Appendix C for a resumé of the Foldy-Wouthuysen representation of the Dirac theory).

In the present context let us first examine how the machinery of Foldy-Wouthuysen transformation technique becomes applicable to the Feshbach-Villars-like form in (2.27). The striking analogy between equation (2.27) and the Dirac equation (2.38) follows from the correspondences: forward propagation of the beam close to the $+z$-direction $\longleftrightarrow$ positive energy Dirac particle, paraxial beam ($|\pi_\perp| \ll \hbar k$) $\longleftrightarrow$ nonrelativistic motion ($|\pi| \ll m_0c$), deviation from paraxial condition (aberrating system) $\longleftrightarrow$ deviation from nonrelativistic situation (relativistic motion). We further note the similarity in the algebraic properties of equation (2.37) and equation (2.44) with σ_z playing the role of β. With the above correspondences we apply the Foldy-Wouthuysen technique to equation (2.27) and obtain a representation in which the strength of the odd operator would be as low as the system under study would require. In this context the measure of the strength of the odd operator $\hat{\mathcal{O}}$

is labeled by the expansion parameter $\frac{1}{k_0}$. It is to be noted that the strength of $\hat{\mathcal{O}}$ in (2.30) is already $\frac{1}{k_0^2}$. This is to be compared with the Dirac case (in Chapter III) where the starting $\hat{\mathcal{O}}$ has the strength $\frac{1}{k_0}$. As the analogy indicates, this procedure would lead us eventually to an expansion of (2.27) into its paraxial part followed by the aberration parts as desired.

Now, following a Foldy-Wouthuysen-like procedure, let us define

$$\begin{pmatrix} \psi_+^{(1)} \\ \psi_-^{(1)} \end{pmatrix} = \mathrm{e}^{\mathrm{i}\hat{S}_1} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}, \qquad \text{with} \quad \hat{S}_1 = \frac{\mathrm{i}}{2}\sigma_z\hat{\mathcal{O}}. \tag{2.45}$$

This transforms equation (2.27) into

$$\begin{aligned} \frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\begin{pmatrix} \psi_+^{(1)} \\ \psi_-^{(1)} \end{pmatrix} &= \left\{\frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\left(\mathrm{e}^{\mathrm{i}\hat{S}_1}\right)\mathrm{e}^{-\mathrm{i}\hat{S}_1} + \mathrm{e}^{\mathrm{i}\hat{S}_1}\hat{\mathrm{H}}\mathrm{e}^{-\mathrm{i}\hat{S}_1}\right\}\begin{pmatrix} \psi_+^{(1)} \\ \psi_-^{(1)} \end{pmatrix} \\ &= \left\{\mathrm{e}^{\mathrm{i}\hat{S}_1}\hat{\mathrm{H}}\mathrm{e}^{-\mathrm{i}\hat{S}_1} - \frac{\mathrm{i}}{k_0}\mathrm{e}^{\mathrm{i}\hat{S}_1}\frac{\partial}{\partial z}\left(\mathrm{e}^{-\mathrm{i}\hat{S}_1}\right)\right\}\begin{pmatrix} \psi_+^{(1)} \\ \psi_-^{(1)} \end{pmatrix} \\ &= \hat{\mathrm{H}}^{(1)}\begin{pmatrix} \psi_+^{(1)} \\ \psi_-^{(1)} \end{pmatrix}, \end{aligned} \tag{2.46}$$

where

$$\hat{\mathrm{H}}^{(1)} = -\sigma_z + \hat{\mathcal{E}}^{(1)} + \hat{\mathcal{O}}^{(1)} \tag{2.47}$$

$$\begin{aligned} \hat{\mathcal{E}}^{(1)} \approx\ & -\frac{q}{\hbar k_0}A_z\mathbb{1} - \left\{\frac{1}{2k_0^2}\left(\hat{D}_\perp^2 - \tilde{k}^2\right) - \frac{1}{8k_0^4}\left(\hat{D}_\perp^2 - \tilde{k}^2\right)^2\right\}\sigma_z \\ & -\frac{1}{32k_0^5}\left\{\left[\left(\hat{D}_\perp^2 - \tilde{k}^2\right), \left[\hat{D}_\perp^2, \frac{q}{\hbar}A_z\right] - \frac{q}{\hbar}\left(\nabla_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z} + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\nabla_\perp\right)\right]\right. \\ & \left. + \left[\hat{D}_\perp^2, \mathrm{i}\frac{\partial}{\partial z}\left(\frac{q^2}{\hbar^2}A_\perp^2 + \tilde{k}^2\right)\right]\right\}\mathbb{1} \end{aligned} \tag{2.48}$$

$$\begin{aligned} \hat{\mathcal{O}}^{(1)} \approx\ & \frac{1}{2k_0^3}\left\{\frac{\mathrm{i}}{\hbar}q\left[\hat{D}_\perp^2, A_z\right] - \frac{\mathrm{i}}{\hbar}q\left(\nabla_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z} + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\nabla_\perp\right)\right. \\ & \left. -\frac{\partial}{\partial z}\left(\frac{q^2}{\hbar^2}A_\perp^2 + \tilde{k}^2\right)\right\}\sigma_y. \end{aligned} \tag{2.49}$$

As expected, the strength of the odd part $\hat{\mathcal{O}}^{(1)}$ in $\hat{\mathrm{H}}^{(1)}$ is now $\frac{1}{k_0^3}$. The next step of transformation of the type in (2.45) with $\hat{\mathcal{O}}$ replaced by $\hat{\mathcal{O}}^{(1)}$ will give an $\hat{\mathrm{H}}^{(2)}$ containing an odd part $\hat{\mathcal{O}}^{(2)}$ of strength $\frac{1}{k_0^4}$. An accuracy up to the order of $\frac{1}{k_0^3}$ is adequate to work out the third order aberrations, to be shown later in this chapter. Hence dropping the odd term $\hat{\mathcal{O}}^{(1)}$, we write

$$\frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\left(\begin{array}{c}\psi_+^{(1)}\\ \psi_-^{(1)}\end{array}\right) = \hat{\mathrm{H}}^{(1)}\left(\begin{array}{c}\psi_+^{(1)}\\ \psi_-^{(1)}\end{array}\right), \tag{2.50}$$

with

$$\hat{\mathrm{H}}^{(1)} = -\sigma_z + \hat{\mathcal{E}}^{(1)} \tag{2.51}$$

Let us now look at $\left(\begin{array}{c}\psi_+^{(1)}\\ \psi_-^{(1)}\end{array}\right)$ corresponding to the plane wave $\psi_{\boldsymbol{k}_0}(\boldsymbol{r})$ in free space. We get

$$\begin{aligned}
\left(\begin{array}{c}\psi^{(1)}_{\boldsymbol{k}_0,+}\\ \psi^{(1)}_{\boldsymbol{k}_0,-}\end{array}\right) &= \mathrm{e}^{\mathrm{i}\hat{S}_1}\left(\begin{array}{c}\psi_{\boldsymbol{k}_0,+}\\ \psi_{\boldsymbol{k}_0,-}\end{array}\right) = \mathrm{e}^{\frac{1}{4k_0^2}\nabla_\perp^2\sigma_x}\left(\begin{array}{c}\psi_{\boldsymbol{k}_0,+}\\ \psi_{\boldsymbol{k}_0,-}\end{array}\right)\\
&\approx \left(\begin{array}{cc}1 & \frac{1}{4k_0^2}\nabla_\perp^2\\ \frac{1}{4k_0^2}\nabla_\perp^2 & 1\end{array}\right)\left(\begin{array}{c}\psi_{\boldsymbol{k}_0,+}\\ \psi_{\boldsymbol{k}_0,-}\end{array}\right)\\
&= \frac{1}{2}\left(\begin{array}{c}1+\frac{k_{0z}}{k_0} - \left(1-\frac{k_{0z}}{k_0}\right)\frac{k_{0\perp}^2}{4k_0^2}\\ 1-\frac{k_{0z}}{k_0} - \left(1+\frac{k_{0z}}{k_0}\right)\frac{k_{0\perp}^2}{4k_0^2}\end{array}\right)\psi_{\boldsymbol{k}_0},
\end{aligned} \tag{2.52}$$

showing that $\psi^{(1)}_{\boldsymbol{k}_0,+} \gg \psi^{(1)}_{\boldsymbol{k}_0,-}$ for a quasiparaxial beam. This result easily extends to the wavepacket of the form in equation(2.35). Thus, in general, we can take $\psi_+^{(1)} \gg \psi_-^{(1)}$ in (2.50) for the beam wavefunctions of interest to us. We can express this property that $\psi_+^{(1)} \gg \psi_-^{(1)}$, or essentially $\psi_-^{(1)} \approx 0$ compared to $\psi_+^{(1)}$, as

$$\sigma_z\left(\begin{array}{c}\psi_+^{(1)}\\ \psi_-^{(1)}\end{array}\right) \approx \mathbb{1}\left(\begin{array}{c}\psi_+^{(1)}\\ \psi_-^{(1)}\end{array}\right). \tag{2.53}$$

With this understanding equation (2.50) is further approximated by

$$\frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\left(\begin{array}{c}\psi_+^{(1)}\\ \psi_-^{(1)}\end{array}\right) = \hat{\mathrm{H}}^{(1)}\left(\begin{array}{c}\psi_+^{(1)}\\ \psi_-^{(1)}\end{array}\right), \tag{2.54}$$

with

$$\begin{aligned}
\hat{\mathrm{H}}^{(1)} \approx &\left(-1 - \frac{q}{\hbar k_0}A_z - \frac{1}{2k_0^2}\left(\hat{D}_\perp^2 - \tilde{k}^2\right)\right.\\
&+\frac{1}{8k_0^4}\left(\hat{D}_\perp^2 - \tilde{k}^2\right)^2 - \frac{1}{32k_0^5}\left\{\left[\left(\hat{D}_\perp^2 - \tilde{k}^2\right),\right.\right.
\end{aligned}$$

$$\left[\hat{D}_\perp^2\,,\,\frac{q}{\hbar}A_z\right] - \frac{q}{\hbar}\left(\nabla_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z} + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\nabla_\perp\right)\Bigg]$$
$$+ \left[\hat{D}_\perp^2\,,\,\mathrm{i}\frac{\partial}{\partial z}\left(\frac{q^2}{\hbar^2}A_\perp^2 + \bar{k}^2\right)\right]\Bigg\}\Bigg)\,\mathbb{1}\,. \tag{2.55}$$

Since we are interested in the z-evolution equation for $\psi(\boldsymbol{r})$ (satisfying (2.22)) we retrace through the following transformations

$$\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = M^{-1}\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix} = M^{-1}\mathrm{e}^{-\mathrm{i}\hat{S}_1}\begin{pmatrix}\psi_+^{(1)}\\ \psi_-^{(1)}\end{pmatrix}\,, \tag{2.56}$$

bearing in mind that $\psi_1 = \psi$ in $\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$. Then we get

$$\begin{aligned}
&\frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}\\
&= \left\{M^{-1}\left[\frac{\mathrm{i}}{k_0}\frac{\partial}{\partial z}\left(\mathrm{e}^{-\mathrm{i}\hat{S}_1}\right)\mathrm{e}^{\mathrm{i}\hat{S}_1} + \mathrm{e}^{-\mathrm{i}\hat{S}_1}\hat{\mathrm{H}}^{(1)}\mathrm{e}^{\mathrm{i}\hat{S}_1}\right]M\right\}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}\\
&= \left\{M^{-1}\left[\mathrm{e}^{-\mathrm{i}\hat{S}_1}\hat{\mathrm{H}}^{(1)}\mathrm{e}^{\mathrm{i}\hat{S}_1} - \frac{\mathrm{i}}{k_0}\mathrm{e}^{-\mathrm{i}\hat{S}_1}\frac{\partial}{\partial z}\left(\mathrm{e}^{\mathrm{i}\hat{S}_1}\right)\right]M\right\}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}\\
&= \hat{\mathrm{H}}_0\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}\,,
\end{aligned} \tag{2.57}$$

where

$$\begin{aligned}
\hat{\mathrm{H}}_0 \approx\; & \left(-1 - \frac{q}{\hbar k_0}A_z - \frac{1}{2k_0^2}\left(\hat{D}_\perp^2 - \bar{k}^2\right)\right)\mathbb{1}\\
& + \frac{1}{4k_0^3}\left(\left[\hat{D}_\perp^2\,,\,\frac{q}{\hbar}A_z\right] - \left\{\frac{q}{\hbar}\left(\nabla_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z}\right.\right.\right.\\
& \left.\left.\left. + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\nabla_\perp\right) - \frac{\mathrm{i}q^2}{\hbar^2}\frac{\partial A_\perp^2}{\partial z}\right\}\right)\sigma_z\\
& + \left(\frac{1}{8k_0^4}\left(\hat{D}_\perp^2 - \bar{k}^2\right)^2 - \frac{1}{16k_0^5}\left\{\left[\left(\hat{D}_\perp^2 - \bar{k}^2\right)\,,\right.\right.\right.\\
& \left[\hat{D}_\perp^2\,,\,\frac{q}{\hbar}A_z\right] - \frac{q}{\hbar}\left(\nabla_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z} + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\nabla_\perp\right)\Bigg]\\
& \left.\left. + \left[\hat{D}_\perp^2\,,\,\mathrm{i}\frac{\partial}{\partial z}\left(\frac{q^2}{\hbar^2}A_\perp^2 + \bar{k}^2\right)\right]\right\}\right)\mathbb{1}\,.
\end{aligned} \tag{2.58}$$

Let us examine the expression for $\hat{\mathrm{H}}_0$ in (2.58). It describes the z-evolution of ψ_1 and ψ_2 independently. We are interested only in the z-evolution of $\psi_1 = \psi$ and hence we can write

$$\frac{\mathrm{i}}{k_0}\frac{\partial\psi}{\partial z} = \hat{H}\psi\,, \tag{2.59}$$

with

$$\begin{aligned}
\hat{H} \;\approx\; & -1 - \frac{q}{\hbar k_0}A_z - \frac{1}{2k_0^2}\left(\hat{D}_\perp^2 - \tilde{k}^2\right) \\
& + \frac{1}{4k_0^3}\left(\left[\hat{D}_\perp^2\,,\,\frac{q}{\hbar}A_z\right] - \left\{\frac{q}{\hbar}\left(\nabla_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z}\right.\right.\right. \\
& \left.\left.\left. + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\nabla_\perp\right) - \frac{\mathrm{i}q^2}{\hbar^2}\frac{\partial A_\perp^2}{\partial z}\right\}\right) \\
& + \frac{1}{8k_0^4}\left(\hat{D}_\perp^2 - \tilde{k}^2\right)^2 - \frac{1}{16k_0^5}\left\{\left[\left(\hat{D}_\perp^2 - \tilde{k}^2\right)\,,\right.\right. \\
& \left.\left[\hat{D}_\perp^2\,,\,\frac{q}{\hbar}A_z\right] - \frac{q}{\hbar}\left(\nabla_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z} + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\nabla_\perp\right)\right] \\
& \left. + \left[\hat{D}_\perp^2\,,\,\mathrm{i}\frac{\partial}{\partial z}\left(\frac{q^2}{\hbar^2}A_\perp^2 + \tilde{k}^2\right)\right]\right\}\,.
\end{aligned} \tag{2.60}$$

Multiplying (2.59) throughout by $\hbar k_0$ we get

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial z} = \hat{\mathcal{H}}_o\psi\,, \tag{2.61}$$

with

$$\begin{aligned}
\hat{\mathcal{H}}_o \;\approx\; & -p_0 - qA_z + \frac{1}{2p_0}\left(\hat{\pi}_\perp^2 + \tilde{p}^2\right) \\
& - \frac{1}{4p_0^2}\left(\left[\hat{\pi}_\perp^2\,,\,qA_z\right] + \left\{\mathrm{i}\hbar q\left(\hat{\boldsymbol{p}}_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z}\right.\right.\right. \\
& \left.\left.\left. + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\hat{\boldsymbol{p}}_\perp\right) - \mathrm{i}\hbar q^2\frac{\partial A_\perp^2}{\partial z}\right\}\right) \\
& + \frac{1}{8p_0^3}\left(\hat{\pi}_\perp^2 + \tilde{p}^2\right)^2 - \frac{1}{16p_0^4}\left\{\left[\left(\hat{\pi}_\perp^2 + \tilde{p}^2\right)\,,\right.\right. \\
& \left.\left[\hat{\pi}_\perp^2\,,\,qA_z\right] + \mathrm{i}q\hbar\left(\hat{\boldsymbol{p}}_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z} + \frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\hat{\boldsymbol{p}}_\perp\right)\right] \\
& \left. - \left[\hat{\pi}_\perp^2\,,\,\mathrm{i}\hbar\frac{\partial}{\partial z}\left(q^2A_\perp^2 + \tilde{p}^2\right)\right]\right\}\,.
\end{aligned} \tag{2.62}$$

Thus, we have obtained $\hat{\mathcal{H}}_o$ which is the required beam optical Hamiltonian corresponding to $\mathrm{i}\hbar\frac{\partial}{\partial z}$ or $-\hat{p}_z$ (= $-$ the canonical momentum in the z-direction as already pointed out in Chapter I). We note that $\hat{\mathcal{H}}_o$ is not hermitian. Physically this is a reflection of the fact that $\int d^2r|\psi(\boldsymbol{r}_\perp, z)|^2$, the probability of finding the particle in the xy-plane at z need not be, in general, a constant along the z-axis; only $\int dz d^2r|\psi(\boldsymbol{r}_\perp, z)|^2$ the total probability for finding the particle somewhere in the entire space ($= 1$) should be conserved at all times considering that the particle cannot just vanish in the absence of any annihilation-creation mechanism. Consequently, the z-evolution of $\psi(\boldsymbol{r}_\perp, z)$, given by (2.61) is not necessarily unitary. First let us write the above beam-optical Hamiltonian as

$$\hat{\mathcal{H}}_o = -p_0 + \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,a} + \hat{\mathcal{H}}_o^{(\lambda_0)} \tag{2.63}$$

where p_0 is the magnitude of the design momentum corresponding to the mean kinetic energy with which a constituent particle of the quasimonoenergetic beam enters the system, from the field-free input region, in a path close to the $+z$ direction, $\hat{\mathbf{H}}_{o,p}$ is the hermitian paraxial Hamiltonian (in general a quadratic expression in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$, $\hat{\mathbf{H}}_{o,a}$ is the hermitian aberration (or perturbation) Hamiltonian (a polynomial of degree > 2 in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$) and $\hat{\mathcal{H}}_o^{(\lambda_0)}$ is a sum of hermitian and antihermitian expressions with explicit λ_0-dependence containing paraxial as well as nonparaxial terms. In the geometrical optics limit ($\lambda_0 \longrightarrow 0$) $\hat{\mathcal{H}}_o^{(\lambda_0)}$ vanishes, unlike $\hat{\mathbf{H}}_{o,p}$ and $\hat{\mathbf{H}}_{o,a}$ which tend to the corresponding classical expressions in this limit. The effect of the nonhermitian terms can be expected to be quite small and negligible. Hence, we can approximate $\hat{\mathcal{H}}_o$, further, to a hermitian $\hat{\mathbf{H}}_o$ by dropping the nonhermitian terms (or, taking $\hat{\mathbf{H}}_o = \frac{1}{2}(\hat{\mathcal{H}}_o + \hat{\mathcal{H}}_o^\dagger)$). Thus we write

$$i\hbar\frac{\partial\psi}{\partial z} = \hat{\mathbf{H}}_o\psi\,, \tag{2.64}$$

with

$$\hat{\mathbf{H}}_o \approx -p_0 - qA_z + \frac{1}{2p_0}\left(\hat{\pi}_\perp^2 + \bar{p}^2\right)$$

$$
\begin{aligned}
&+\frac{1}{8p_0^3}\left(\hat{\pi}_\perp^2+\tilde{p}^2\right)^2-\frac{1}{16p_0^4}\left\{\left[\left(\hat{\pi}_\perp^2+\tilde{p}^2\right),\right.\right.\\
&\left.\left[\hat{\pi}_\perp^2, qA_z\right]+\mathrm{i}q\hbar\left(\hat{\boldsymbol{p}}_\perp\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial z}+\frac{\partial \boldsymbol{A}_\perp}{\partial z}\cdot\hat{\boldsymbol{p}}_\perp\right)\right]\\
&\left.-\left[\hat{\pi}_\perp^2, \mathrm{i}\hbar\frac{\partial}{\partial z}\left(q^2A_\perp^2+\tilde{p}^2\right)\right]\right\}. \qquad (2.65)
\end{aligned}
$$

Having obtained the required basic beam-optical Hamiltonian operator $\hat{\mathcal{H}}_o$ we can now proceed to get the desired relation between $\psi_{\rm in}$ and $\psi_{\rm out}$ using the well-known techniques of quantum mechanics.

The formal integration of (2.61) gives, as is well-known, for any pair of points $\left\{\left(z^{(1)}, z^{(2)}\right)\middle| z^{(2)}>z^{(1)}\right\}$ on the z-axis,

$$
\left|\psi\left(z^{(2)}\right)\right\rangle=\hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}\right)\left|\psi\left(z^{(1)}\right)\right\rangle, \qquad (2.66)
$$

with

$$
\mathrm{i}\hbar\frac{\partial}{\partial z}\hat{\mathcal{T}}\left(z, z^{(1)}\right)=\hat{\mathcal{H}}_o\hat{\mathcal{T}}\left(z, z^{(1)}\right), \quad \hat{\mathcal{T}}\left(z^{(1)}, z^{(1)}\right)=\hat{\mathcal{I}}
$$

$$
\begin{aligned}
\hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}\right) &= \wp\left\{\exp\left(-\frac{\mathrm{i}}{\hbar}\int_{z^{(1)}}^{z^{(2)}} dz\,\hat{\mathcal{H}}_o(z)\right)\right\}\\
&= \hat{\mathcal{I}}-\frac{\mathrm{i}}{\hbar}\int_{z^{(1)}}^{z^{(2)}} dz\,\hat{\mathcal{H}}_o(z)\\
&\quad+\left(-\frac{\mathrm{i}}{\hbar}\right)^2\int_{z^{(1)}}^{z^{(2)}} dz\int_{z^{(1)}}^{z} dz'\,\hat{\mathcal{H}}_o(z)\hat{\mathcal{H}}_o(z')\\
&\quad+\left(-\frac{\mathrm{i}}{\hbar}\right)^3\int_{z^{(1)}}^{z^{(2)}} dz\int_{z^{(1)}}^{z} dz'\int_{z^{(1)}}^{z'} dz''\,\hat{\mathcal{H}}_o(z)\hat{\mathcal{H}}_o(z')\hat{\mathcal{H}}_o(z'')\\
&\quad+\ldots, \qquad (2.67)
\end{aligned}
$$

where $\hat{\mathcal{I}}$ is the identity operator and $\wp$ denotes the path-ordered exponential. Except in a very few cases there is no closed form expression for $\hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}\right)$. For our purpose the most convenient form of the expression for the z-evolution operator $\hat{\mathcal{T}}(z^{(2)}, z^{(1)})$, or the z-propagator, is

$$
\hat{\mathcal{T}}(z^{(2)}, z^{(1)})=\exp\left\{-\frac{\mathrm{i}}{\hbar}\hat{T}(z^{(2)}, z^{(1)})\right\}, \qquad (2.68)
$$

with

$$
\begin{aligned}
\hat{T}(z^{(2)}, z^{(1)}) & \\
= & \int_{z^{(1)}}^{z^{(2)}} dz\, \hat{\mathcal{H}}_o(z) \\
& + \frac{1}{2}\left(-\frac{\mathrm{i}}{\hbar}\right) \int_{z^{(1)}}^{z^{(2)}} dz \int_{z^{(1)}}^{z} dz' \left[\hat{\mathcal{H}}_o(z)\,,\, \hat{\mathcal{H}}_o(z')\right] \\
& + \frac{1}{6}\left(-\frac{\mathrm{i}}{\hbar}\right)^2 \int_{z^{(1)}}^{z^{(2)}} dz \int_{z^{(1)}}^{z} dz' \int_{z^{(1)}}^{z'} dz'' \\
& \qquad \left\{\left[\left[\hat{\mathcal{H}}_o(z)\,,\, \hat{\mathcal{H}}_o(z')\right]\,,\, \hat{\mathcal{H}}_o(z'')\right] \right. \\
& \qquad \left. + \left[\left[\hat{\mathcal{H}}_o(z'')\,,\, \hat{\mathcal{H}}_o(z')\right]\,,\, \hat{\mathcal{H}}_o(z)\right]\right\} \\
& + \ldots\,,
\end{aligned}
\tag{2.69}
$$

as given by the Magnus formula (see Appendix A for details).

It is to be noted that when $\hat{\mathcal{H}}_o$ is hermitian (or when approximated to the hermitian $\hat{\mathbf{H}}_o$) $\hat{T}$ becomes hermitian and consequently $\hat{\mathcal{T}}$ is unitary. In such a case $\hat{T}$ and $\hat{\mathcal{T}}$ will be denoted by $\hat{\mathbb{T}}$ and $\hat{U}$ respectively. Then, equation (2.64) leads to

$$
\left|\psi\left(z^{(2)}\right)\right\rangle = \hat{U}\left(z^{(2)}, z^{(1)}\right)\left|\psi\left(z^{(1)}\right)\right\rangle\,, \tag{2.70}
$$

where

$$
\hat{U}\left(z^{(2)}, z^{(1)}\right) = \exp\left\{-\frac{\mathrm{i}}{\hbar}\hat{\mathbb{T}}\left(z^{(2)}, z^{(1)}\right)\right\}, \tag{2.71}
$$

with

$$
\begin{aligned}
\hat{\mathbb{T}}\left(z^{(2)}, z^{(1)}\right) & \\
= & \int_{z^{(1)}}^{z^{(2)}} dz\, \hat{\mathbf{H}}_o(z) \\
& + \frac{1}{2}\left(-\frac{\mathrm{i}}{\hbar}\right) \int_{z^{(1)}}^{z^{(2)}} dz \int_{z^{(1)}}^{z} dz' \left[\hat{\mathbf{H}}_o(z)\,,\, \hat{\mathbf{H}}_o(z')\right] \\
& + \frac{1}{6}\left(-\frac{\mathrm{i}}{\hbar}\right)^2 \int_{z^{(1)}}^{z^{(2)}} dz \int_{z^{(1)}}^{z} dz' \int_{z^{(1)}}^{z'} dz'' \\
& \qquad \left\{\left[\left[\hat{\mathbf{H}}_o(z)\,,\, \hat{\mathbf{H}}_o(z')\right]\,,\, \hat{\mathbf{H}}_o(z'')\right] \right. \\
& \qquad \left. + \left[\left[\hat{\mathbf{H}}_o(z'')\,,\, \hat{\mathbf{H}}_o(z')\right]\,,\, \hat{\mathbf{H}}_o(z)\right]\right\} \\
& + \ldots\,.
\end{aligned}
\tag{2.72}
$$

In order to understand the electron optical image formation we should work with the coordinate representation in the Schrödinger picture. So, we write (2.66), the integral form of the optical Schrödinger equation (2.61), as

$$\psi\left(\boldsymbol{r}_\perp^{(2)}, z^{(2)}\right) = \int d^2r^{(1)} G\left(\boldsymbol{r}_\perp^{(2)}, z^{(2)}; \boldsymbol{r}_\perp^{(1)}, z^{(1)}\right) \psi\left(\boldsymbol{r}_\perp^{(1)}, z^{(1)}\right) , \tag{2.73}$$

where the Green's function $G\left(\boldsymbol{r}_\perp^{(2)}, z^{(2)}; \boldsymbol{r}_\perp^{(1)}, z^{(1)}\right)$ is given by

$$\begin{aligned} & G\left(\boldsymbol{r}_\perp^{(2)}, z^{(2)}; \boldsymbol{r}_\perp^{(1)}, z^{(1)}\right) \\ & = \left\langle \boldsymbol{r}_\perp^{(2)} \mid \hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}\right) \mid \boldsymbol{r}_\perp^{(1)} \right\rangle \\ & = \int d^2r \delta^{2*}\left(\boldsymbol{r}_\perp - \boldsymbol{r}_\perp^{(2)}\right) \hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}\right) \delta^2\left(\boldsymbol{r}_\perp - \boldsymbol{r}_\perp^{(1)}\right) . \end{aligned} \tag{2.74}$$

This is only a formal expression for the Green's function as the matrix element of $\hat{\mathcal{T}}$ and its computation is, generally, quite a difficult task beyond the case of paraxial approximation.

When we want to relate the values of the quantum averages of the observables of the beam at two different points along the axis of the system we can use the Heisenberg picture. The quantum average, or the expectation value, of any observable, say O, associated with the hermitian operator $\hat{O}$ is given as follows: for the state $|\psi(z)\rangle$ at the xy-plane at the point z

$$\langle O\rangle(z) = \frac{\langle\psi(z)|\hat{O}|\psi(z)\rangle}{\langle\psi(z)|\psi(z)\rangle} , \tag{2.75}$$

with the notations

$$\begin{aligned} \langle\psi(z)|\hat{O}|\psi(z)\rangle &= \int d^2r \psi^*(\boldsymbol{r}_\perp, z)\hat{O}\psi(\boldsymbol{r}_\perp, z) \qquad (2.76) \\ \langle\psi(z)|\psi(z)\rangle &= \int d^2r \psi^*(\boldsymbol{r}_\perp, z)\psi(\boldsymbol{r}_\perp, z) . \qquad (2.77) \end{aligned}$$

Some times we shall denote $\langle O\rangle(z)$ by $\langle\hat{O}\rangle(z)$ also. Now, in view of the relation in (2.66), we have

$$\langle O\rangle\left(z^{(2)}\right)$$

$$
\begin{aligned}
&= \frac{\left\langle \psi\left(z^{(2)}\right) \middle| \hat{O} \middle| \psi\left(z^{(2)}\right) \right\rangle}{\left\langle \psi\left(z^{(2)}\right) \middle| \psi\left(z^{(2)}\right) \right\rangle} \\
&= \frac{\left\langle \psi\left(z^{(1)}\right) \middle| \hat{\mathcal{T}}^{\dagger}\left(z^{(2)}, z^{(1)}\right) \hat{O}\, \hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}\right) \middle| \psi\left(z^{(1)}\right) \right\rangle}{\left\langle \psi\left(z^{(1)}\right) \middle| \hat{\mathcal{T}}^{\dagger}\left(z^{(2)}, z^{(1)}\right) \hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}\right) \middle| \psi\left(z^{(1)}\right) \right\rangle}, \qquad (2.78)
\end{aligned}
$$

leading to the required transfer map giving the expectation values of the observables in the plane at $z^{(2)}$ in terms of their values in the plane at $z^{(1)}$. It should be noted that $\langle O \rangle \left(z^{(2)}\right)$ is real even if the transfer operator $\hat{\mathcal{T}}$ is nonunitary. We will be using equations (2.73) and (2.74) to understand diffraction in the field-free space and electron optical image formation using the round magnetic lens. Equation (2.78) will be the basis for our understanding of the focusing properties of electron lenses. Later, we shall see how this formalism, developed for the scalar wavefunction so far, gets generalized to the case of the Dirac spinor wavefunction. In the classical limit, equation (2.78) leads to the Lie algebraic treatment of geometrical charged-particle optics pioneered by Dragt *et al.* (*e.g.*, see ([15]-[19], and [20, 21]). The differential form of (2.78) corresponds to the Heisenberg equation of motion for the observables and would lead to the trajectory equations of geometrical optics in the classical limit in accordance with the correspondence principle (Ehrenfest's theorem).

Finally, it may be noted that the formalism developed here should be suitable for use in light optics also in certain situations, in the theory of graded index fibres, for examples. In fact, wherever one studies single-frequency propagating rays making small angles with the principal direction of propagation the above method could be used: in ocean acoustics, for example (see *e.g.*, [35] and references therein).

2.2 Applications

2.2.1 Free Propagation: Diffraction

Let us first apply the above general formalism to the case of a monoenergetic quasi-paraxial beam of particles moving in free space ($\phi(\boldsymbol{r}) = 0$, $\boldsymbol{A} = (0, 0, 0)$). Now, our system is an infinite slab of free space situated perpendicular to the z-axis between the coordinates $z_{\rm in}$ and $z_{\rm out}$. From (2.65) the corresponding optical Hamiltonian is is read as

$$\hat{\mathcal{H}}_o = \hat{\mathbf{H}}_o \approx -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2 + \frac{1}{8p_0^3}\hat{p}_\perp^4 \,, \tag{2.79}$$

which in the paraxial approximation simplifies further to

$$\hat{\mathbf{H}}_o \approx \hat{\mathbf{H}}_{o,p} = -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2 \,. \tag{2.80}$$

Taking, $z^{(1)} = z_{\rm in}$ and $z^{(2)} = z_{\rm out}$, equation (2.70) becomes

$$|\psi_{\rm out}\rangle = \hat{U}_{D,p}\left(z_{\rm out}, z_{\rm in}\right)|\psi_{\rm in}\rangle \,, \tag{2.81}$$

with

$$\begin{aligned} \hat{U}_{D,p}\left(z_{\rm out}, z_{\rm in}\right) &= \exp\left\{-\frac{\rm i}{\hbar}\hat{\mathbb{T}}_{D,p}\left(z_{\rm out}, z_{\rm in}\right)\right\} \\ &= \exp\left\{-\frac{\rm i}{\hbar}\Delta z \hat{\mathbf{H}}_{o,p}\right\} \\ &= \exp\left\{-\frac{\rm i}{\hbar}\Delta z\left(-p_0 + \frac{1}{2p_0}\hat{p}_\perp^2\right)\right\} \,, \\ & \qquad \Delta z = z_{\rm out} - z_{\rm in} \,, \end{aligned} \tag{2.82}$$

where the subscripts p and D indicate, respectively, paraxial approximation and drift in free space. In coordinate representation

$$\psi\left(\boldsymbol{r}_{\perp,{\rm out}}, z_{\rm out}\right) = \int d^2 r_{\rm in}\, G_{D,p}\left(\boldsymbol{r}_{\perp,{\rm out}}, z_{\rm out}; \boldsymbol{r}_{\perp,{\rm in}}, z_{\rm in}\right)\psi\left(\boldsymbol{r}_{\perp,{\rm in}}, z_{\rm in}\right) \,, \tag{2.83}$$

where

$$G_{D,p}\left(\boldsymbol{r}_{\perp,{\rm out}}, z_{\rm out}; \boldsymbol{r}_{\perp,{\rm in}}, z_{\rm in}\right)$$

$$= \left\langle \boldsymbol{r}_{\perp,\text{out}} \left| \exp\left\{-\frac{\mathrm{i}}{\hbar}\Delta z\left(-p_0 + \frac{1}{2p_0}\hat{p}_\perp^2\right)\right\} \right| \boldsymbol{r}_{\perp,\text{in}} \right\rangle$$
$$= \mathrm{e}^{\mathrm{i}p_0\Delta z/\hbar} \left\langle \boldsymbol{r}_{\perp,\text{out}} \left| \exp\left\{-\frac{\mathrm{i}}{\hbar}\Delta z\left(\frac{1}{2p_0}\hat{p}_\perp^2\right)\right\} \right| \boldsymbol{r}_{\perp,\text{in}} \right\rangle . \qquad (2.84)$$

The matrix element $\left\langle \boldsymbol{r}_{\perp,\text{out}} \left| \exp\left\{-\frac{\mathrm{i}}{\hbar}\Delta z\left(\frac{1}{2p_0}\hat{p}_\perp^2\right)\right\} \right| \boldsymbol{r}_{\perp,\text{in}} \right\rangle$ is readily calculated and is just the well-known Green's function of a nonrelativistic free particle of mass p_0 moving in the xy-plane and corresponding to a time interval Δz. (see Appendix D for details of the calculation). Then, we have

$$G_{D,p}\left(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}; \boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}\right)$$
$$= \mathrm{e}^{\mathrm{i}p_0\Delta z/\hbar}\left(\frac{p_0}{2\pi\mathrm{i}\hbar\Delta z}\right)\exp\left\{\frac{\mathrm{i}p_0}{2\hbar\Delta z}\left|\left(\boldsymbol{r}_{\perp,\text{out}} - \boldsymbol{r}_{\perp,\text{in}}\right)\right|^2\right\} . \qquad (2.85)$$

Now, equation (2.83) becomes

$$\psi\left(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}\right)$$
$$= \mathrm{e}^{\mathrm{i}p_0\Delta z/\hbar}\left(\frac{p_0}{2\pi\mathrm{i}\hbar\Delta z}\right)\int d^2r_{\text{in}}$$
$$\exp\left\{\frac{\mathrm{i}p_0}{2\hbar\Delta z}\left|\left(\boldsymbol{r}_{\perp,\text{out}} - \boldsymbol{r}_{\perp,\text{in}}\right)\right|^2\right\}\psi\left(\boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}\right)$$
$$= \frac{1}{\mathrm{i}\lambda_0\Delta z}\mathrm{e}^{\mathrm{i}k_0\Delta z}\int\int dx_{\text{in}}\, dy_{\text{in}}$$
$$\exp\left\{\frac{\mathrm{i}k_0}{2\Delta z}\left[\left(x_{\text{out}} - x_{\text{in}}\right)^2 + \left(y_{\text{out}} - y_{\text{in}}\right)^2\right]\right\}\psi\left(\boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}\right) \qquad (2.86)$$

which is the well-known Fresnel diffraction formula; λ_0 is the de Broglie wavelength $2\pi\hbar/p_0$. Here, z_{in}-plane (the xy-plane at $z = z_{\text{in}}$) is the plane of the diffracting object and z_{out}-plane (the xy-plane at $z = z_{\text{out}}$) is the observation plane; $\psi\left(\boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}\right)$ is the wavefunction on the exit side of the diffracting object and $\left|\psi\left(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}\right)\right|^2$ gives the intensity distribution of the diffraction pattern at the observation plane. It is clear that the approximation of $\hat{\mathbf{H}}_o$ as in (2.80), dropping terms of order higher than second in $\hat{\boldsymbol{p}}_\perp$, essentially corresponds to the traditional approximation used in deriving the Fresnel diffraction formula from the general Kirchhoff's result. We can also recognize now that equations (2.66)–(2.69), (2.73) and (2.74), along with (2.62), represent, in operator form, the general theory of charged-particle diffraction in presence of electromagnetic fields (for more details on diffraction theory, see [5] Chapters 59 and 60).

Actually, the free propagation case can be treated exactly as noted already in the classical theory. The expression for the Hamiltonian $\hat{\mathbf{H}}_o$ given above, in (2.79), is an approximation for the exact result, $\hat{\mathbf{H}}_o = -\sqrt{p_0^2 - \hat{p}_\perp^2}$, obtained by quantizing the classical expression for $-p_z$ for a particle of momentum p_0; this exact result will be obtained in the infinite series form in our approach also if we continue the Foldy-Wouthuysen transformation process up to all orders. Hence the exact form of the Green's function is given by the matrix element of $\exp\left(\frac{i}{\hbar}\Delta z\sqrt{p_0^2 - \hat{p}_\perp^2}\right)$ and it can be shown that an explicit evaluation of this matrix element is possible leading to the well-known exact scalar wave Green's function for the plane (see [15]).

Let us now work out the transfer maps for the expectation values of the transverse coordinates $\langle \boldsymbol{r}_\perp \rangle$ and their conjugate momenta $\langle \boldsymbol{p}_\perp \rangle$ in the case of free propagation. Let a particle of the input beam be associated with a wavepacket $\psi(\boldsymbol{r}_\perp, z)$ having $\langle \boldsymbol{r}_\perp \rangle_{\rm in} = \langle \boldsymbol{r}_\perp \rangle (z_{\rm in})$ and $\langle \boldsymbol{p}_\perp \rangle_{\rm in} = \langle \boldsymbol{p}_\perp \rangle (z_{\rm in})$ as the average values of the transverse coordinates and momenta at the $z_{\rm in}$-plane. In the geometrical optics picture the particle corresponds to a ray intersecting the $z_{\rm in}$-plane at the point $\boldsymbol{r}_\perp = \langle \boldsymbol{r}_\perp \rangle_{\rm in}$ with the gradient $\frac{d\boldsymbol{r}_\perp}{dz} = \boldsymbol{p}_\perp / p_z \approx \boldsymbol{p}_\perp / p_0 = \langle \boldsymbol{p}_\perp \rangle_{\rm in} / p_0$; paraxiality condition is seen to be $\frac{dx}{dz} \ll 1$, $\frac{dy}{dz} \ll 1$. From the formula of (2.78), and using (2.71) and (2.72), we get

$$
\begin{aligned}
\langle \boldsymbol{r}_\perp \rangle_{\rm out} &= \langle \boldsymbol{r}_\perp \rangle (z_{\rm out}) \\
&= \left\langle \psi_{\rm in} \left| e^{\frac{i}{\hbar}\Delta z \hat{\mathrm{H}}_o} \boldsymbol{r}_\perp e^{-\frac{i}{\hbar}\Delta z \hat{\mathrm{H}}_o} \right| \psi_{\rm in} \right\rangle \qquad (2.87) \\
\langle \boldsymbol{p}_\perp \rangle_{\rm out} &= \langle \boldsymbol{p}_\perp \rangle (z_{\rm out}) \\
&= \left\langle \psi_{\rm in} \left| e^{\frac{i}{\hbar}\Delta z \hat{\mathrm{H}}_o} \hat{\boldsymbol{p}}_\perp e^{-\frac{i}{\hbar}\Delta z \hat{\mathrm{H}}_o} \right| \psi_{\rm in} \right\rangle , \qquad (2.88)
\end{aligned}
$$

with $\hat{\mathbf{H}}_o$ as given in (2.79). Using the relation

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]] + \frac{1}{3!}[\hat{A},[\hat{A},[\hat{A},\hat{B}]]] + \ldots \tag{2.89}$$

valid for any pair of operators $(\hat{A},\hat{B})$, we have from equations (2.87) and (2.88)

$$\begin{aligned}
\langle \boldsymbol{r}_\perp \rangle_{\rm out} &= \langle \boldsymbol{r}_\perp \rangle_{\rm in} + \Delta z \left\langle \frac{\hat{\boldsymbol{p}}_\perp}{p_0}\left(1 + \frac{1}{2p_0^2}p_\perp^2\right)\right\rangle_{\rm in} \\
&\approx \langle \boldsymbol{r}_\perp \rangle_{\rm in} + \Delta z \left\langle \frac{\hat{\boldsymbol{p}}_\perp}{\sqrt{p_0^2 - p_\perp^2}}\right\rangle_{\rm in} \\
&= \langle \boldsymbol{r}_\perp \rangle_{\rm in} + \Delta z \left\langle \frac{\hat{\boldsymbol{p}}_\perp}{p_z}\right\rangle_{\rm in} \\
&= \langle \boldsymbol{r}_\perp \rangle_{\rm in} + \Delta z \left\langle \frac{d\boldsymbol{r}_\perp}{dz}\right\rangle_{\rm in}
\end{aligned} \tag{2.90}$$

$$\langle \boldsymbol{p}_\perp \rangle_{\rm out} = \langle \boldsymbol{p}_\perp \rangle_{\rm in}\,, \quad \text{or} \quad \left\langle \frac{d\boldsymbol{r}_\perp}{dz}\right\rangle_{\rm out} = \left\langle \frac{d\boldsymbol{r}_\perp}{dz}\right\rangle_{\rm in}, \tag{2.91}$$

confirming the rectilinear propagation law for the free ray. In matrix form, we have

$$\begin{pmatrix} \langle \boldsymbol{r}_\perp \rangle \\ \left\langle \frac{d\boldsymbol{r}_\perp}{dz} \right\rangle \end{pmatrix}_{\rm out} = \begin{pmatrix} 1 & \Delta z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \langle \boldsymbol{r}_\perp \rangle \\ \left\langle \frac{d\boldsymbol{r}_\perp}{dz} \right\rangle \end{pmatrix}_{\rm in}, \tag{2.92}$$

giving the familiar transfer matrix for free propagation in terms of the traditional ray variables $\left(\boldsymbol{r}_\perp, \frac{d\boldsymbol{r}_\perp}{dz}\right)$. Taking $p_z \approx p_0$, we can write

$$\begin{pmatrix} \langle \boldsymbol{r}_\perp \rangle \\ \langle \boldsymbol{p}_\perp / p_0 \rangle \end{pmatrix}_{\rm out} = \begin{pmatrix} 1 & \Delta z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \langle \boldsymbol{r}_\perp \rangle \\ \langle \boldsymbol{p}_\perp / p_0 \rangle \end{pmatrix}_{\rm in}, \tag{2.93}$$

giving the transfer matrix in terms of the canonical phase-space variables $(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp)$ (see [36, 37], for a treatment of geometrical electron optics using the canonical variables $(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp)$).

2.2.2 Axially Symmetric Magnetic Lens: Electron Optical Imaging

The axially symmetric magnetic lens, or the round magnetic lens, is the central part of any electron microscope. In practice, the round magnetic lenses of electron microscopy are convergent lenses. The round magnetic lens is an axially symmetric magnetic field completely characterized by the potentials

$$\begin{aligned} \phi(\boldsymbol{r}) &= 0 \qquad (2.94) \\ \boldsymbol{A} &= \left(-\frac{y}{2}\Pi(\boldsymbol{r}_\perp, z)\,,\, \frac{x}{2}\Pi(\boldsymbol{r}_\perp, z)\,,\, 0\right)\,, \qquad (2.95) \end{aligned}$$

with

$$\begin{aligned} \Pi(\boldsymbol{r}_\perp, z) &= \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!}\left(-\frac{r_\perp^2}{4}\right)^n B^{(2n)}(z) \\ &= B(z) - \frac{1}{8}r_\perp^2 B''(z) + \frac{1}{192}r_\perp^4 B''''(z) - \ldots\,, \end{aligned} \qquad (2.96)$$

The corresponding magnetic field is

$$\begin{aligned} \boldsymbol{B}_\perp &= -\frac{1}{2}\left(B'(z) - \frac{1}{8}r_\perp^2 B'''(z) + \ldots\right)\boldsymbol{r}_\perp \\ B_z &= B(z) - \frac{1}{4}r_\perp^2 B''(z) + \frac{1}{64}r_\perp^4 B''''(z) - \ldots\,, \end{aligned} \qquad (2.97)$$

Due to the axial symmetry of the system the potential has only terms of odd order in $\boldsymbol{r}_\perp$. The linear terms in the potential govern the paraxial behaviour, and the higher ones give rise to aberrations of the corresponding order. So we retain terms up to third order in the potentials to get the third order aberrations. The procedure which we will follow can be used to compute aberrations of any higher order. Thus, the potential is approximated by

$$\begin{aligned} \phi(\boldsymbol{r}) &= 0 \qquad (2.98) \\ \boldsymbol{A} &= \left(-\frac{1}{2}y\Pi(\boldsymbol{r}_\perp, z)\,,\, \frac{1}{2}x\Pi(\boldsymbol{r}_\perp, z)\,,\, 0\right)\,, \\ & \qquad \text{with } \Pi(\boldsymbol{r}_\perp, z) = B(z) - \frac{1}{8}r_\perp^2 B''(z)\,. \qquad (2.99) \end{aligned}$$

Then, the beam-optical Hamiltonian (2.62) becomes

$$\hat{\mathcal{H}}_o = \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,(4)} + \hat{\mathcal{H}}_o^{(\lambda_0)} \qquad (2.100)$$

$$
\begin{aligned}
\hat{\mathbf{H}}_{o,p} &= = -p_0 + \frac{1}{2p_0}\left(\hat{p}_\perp^2 + \frac{1}{4}q^2B^2(z)r_\perp^2 - qB(z)\hat{L}_z\right), && (2.101)\\
\hat{\mathbf{H}}_{o,(4)} &= \frac{1}{8p_0^3}\hat{p}_\perp^4 - \frac{1}{2p_0^2}\alpha\hat{p}_\perp^2\hat{L}_z - \frac{1}{8p_0}\alpha^2\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp + \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)^2 \\
&\quad + \frac{3}{8p_0}\alpha^2\left(\hat{p}_\perp^2 r_\perp^2 + r_\perp^2\hat{p}_\perp^2\right) + \frac{1}{8}\left(\alpha'' - 4\alpha^3\right)\hat{L}_z r_\perp^2 \\
&\quad + \frac{p_0}{8}\left(\alpha^4 - \alpha\alpha''\right)r_\perp^4, \\
&\qquad \text{with } \alpha = \frac{qB(z)}{2p_0}, \quad \hat{L}_z = x\hat{p}_y - y\hat{p}_x && (2.102)\\
\hat{\mathcal{H}}_o^{(\lambda_0)} &= \lambda_0 - \text{dependent constant } + \hat{\mathbf{H}}_{o,p}^{(\lambda_0)} + \hat{\mathbf{H}}_{o,(4)}^{(\lambda_0)} + \mathcal{A}_{o,p}^{(\lambda_0)} && (2.103)\\
\hat{\mathbf{H}}_{o,p}^{(\lambda_0)} &= \frac{\lambda_0^2 q^2}{64\pi^2 p_0^2}B(z)B'(z)\left(\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp + \hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp\right) && (2.104)\\
\hat{\mathbf{H}}_{o,(4)}^{(\lambda_0)} &= \frac{\lambda_0^2 q}{256\pi^2 p_0^2}B'''(z)\hat{L}_z\left(\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp + \hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp\right) && (2.105)\\
&\quad - \frac{\lambda_0^2 q^2}{512\pi^2 p_0^2}\left(B(z)B''(z)\right)'\left\{r_\perp^2, \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp + \hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp\right\}, \\
\mathcal{A}_{o,p}^{(\lambda_0)} &= \frac{\mathrm{i}\lambda_0}{8\pi p_0}\left(\frac{1}{2}q^2B(z)B'(z)r_\perp^2 - qB'(z)\hat{L}_z\right). && (2.106)
\end{aligned}
$$

The reason for partitioning $\hat{\mathcal{H}}_o$ in the above manner will be clear as we proceed. It is to be noted that p_0 is the magnitude of the design momentum corresponding to the mean kinetic energy with which a constituent particle of the quasimonoenergetic beam enters the system, from the field-free input region, in a path close to the $+z$ direction, $\hat{\mathbf{H}}_{o,p}$ is the hermitian paraxial Hamiltonian (a quadratic expression in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$, $\hat{\mathbf{H}}_{o,(4)}$ is the hermitian aberration (or perturbation) Hamiltonian (a polynomial of degree 4 in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$) and $\hat{\mathcal{H}}_o^{(\lambda_0)}$ is a sum of hermitian and antihermitian expressions with explicit λ_0-dependence containing paraxial as well as nonparaxial terms. In the geometrical optics limit ($\lambda_0 \longrightarrow 0$) $\hat{\mathcal{H}}_o^{(\lambda_0)}$ vanishes, unlike $\hat{\mathbf{H}}_{o,p}$ and $\hat{\mathbf{H}}_{o,(4)}$ which tend to the corresponding classical expressions in this limit. The above Hamiltonian is to be compared with the classical counterpart discussed in Chapter I. The effect of the contributions from $\hat{\mathcal{H}}_o^{(\lambda_0)}$ will be noted in separately.

To summarize, we have obtained the beam-optical Hamiltonian in the quartic approximation, adequate to describe the third order aberrations and also demonstrate the quantum contributions. To proceed further, we assume the lens to lie between the two planes bounded by $z = z_l$ and $z = z_r$. That is, we assume the magnetic fields outside the lens to be negligible and consequently $B(z < z_l) \approx 0$ and $B(z > z_r) \approx 0$. We shall be examining the system sequentially, first the paraxial part, then the aberrating part and finally the effect of the explicit λ_0-dependent contributions.

Note that $\hat{L}_z$, the z-component of the angular momentum $(\boldsymbol{r} \times \hat{\boldsymbol{p}})$, commutes with the total Hamiltonian $\hat{\mathcal{H}}_o$ as a consequence of the axial symmetry of the system. So we further partition the beam-optical Hamiltonian $\hat{\mathcal{H}}_o$ by rewriting the paraxial part. Introducing the notations

$$\begin{aligned} \theta'(z) &= \frac{qB(z)}{2p_0} && (2.107) \\ F(z) &= \alpha^2 = \frac{q^2B^2(z)}{4p_0^2} && (2.108) \\ \underline{\hat{\mathbf{H}}}_{o,p} &= \frac{1}{2p_0}\hat{p}_\perp^2 + \frac{1}{2}p_0F(z)r_\perp^2\,, && (2.109) \end{aligned}$$

let us write the Hamiltonian $\hat{\mathbf{H}}_{o,p}$ in (2.101) as

$$\hat{\mathbf{H}}_{o,p} = -p_0 - \theta'(z)\hat{L}_z + \underline{\hat{\mathbf{H}}}_{o,p}\,. \qquad (2.110)$$

As expected, this will enable us to write the paraxial transfer operator as a product of three transfer operators, as is done in the classical case, by writing the transfer matrix as a product of simpler transfer matrices. The Hamiltonian $\underline{\hat{\mathbf{H}}}_{o,p}$ is like the Hamiltonian of an isotropic two-dimensional harmonic oscillator (in the xy-plane) with z playing the role of time and p_0 playing the role of the oscillator mass. And, the corresponding z-dependent frequency is $\sqrt{F(z)}$. Thus, apart from the two terms, the constant p_0 responsible for a phase factor and the second term $\left(-\theta'(z)\hat{L}_z\right)$ responsible for image rotation, the beam-optical Hamiltonian governing the evolution of the beam along the z-axis is like the Hamiltonian of a two-dimensional harmonic oscillator with time-dependent frequency. This analogy between paraxial beam optics and harmonic oscillator is well-known (*e.g.*, see [38] in the context of light optics and [39] in the context of charged-particle optics). Since $\hat{\mathbf{H}}_{o,p}$ is hermitian the corresponding transfer operator will be unitary as in equations (2.71) and (2.72).

Let us now consider the transfer operator $\hat{U}_p(z, z_o)$ for a general $z > z_o$ where z_o is the coordinate of the object plane. Later we shall determine it specifically at the z_i-plane, where the image is formed, and the corresponding $\hat{U}_p(z_i, z_o)$. Since p_0, $\hat{L}_z$ and $\hat{\underline{\mathbf{H}}}_{o,p}$ commute with each other, we can write

$$\hat{U}_p(z, z_o) = \mathrm{e}^{\frac{\mathrm{i}}{\hbar} p_0 (z - z_o)} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\theta(z, z_o)\hat{L}_z} \hat{\underline{U}}_p(z, z_o)\,, \tag{2.111}$$

where

$$\theta(z, z_o) = \int_{z_o}^{z} dz\, \theta'(z) \tag{2.112}$$

and

$$\hat{\underline{U}}_p(z, z_o) = \exp\left(-\frac{\mathrm{i}}{\hbar} \hat{\underline{\mathbb{T}}}_p(z, z_o)\right)\,, \tag{2.113}$$

the transfer operator corresponding to the Hamiltonian $\hat{\underline{\mathbf{H}}}_{o,p}$, is to be computed using (2.72) with $\hat{\mathbf{H}}_o$ replaced by $\hat{\underline{\mathbf{H}}}_{o,p}$. Then,

$$\psi(\boldsymbol{r}_\perp, z) = \int d^2 r_o G_p(\boldsymbol{r}_\perp, z; \boldsymbol{r}_{\perp,o}, z_o)\, \psi(\boldsymbol{r}_{\perp,o}, z_o) \tag{2.114}$$

with the Green's function

$$\begin{aligned} G_p(\boldsymbol{r}_\perp, z; \boldsymbol{r}_{\perp,o}, z_o) &= \left\langle \boldsymbol{r}_\perp \left| \hat{U}_p(z, z_o) \right| \boldsymbol{r}_{\perp,o} \right\rangle \\ &= \mathrm{e}^{\frac{\mathrm{i}}{\hbar} p_0 (z - z_o)} \int d^2\bar{r} \left\langle \boldsymbol{r}_\perp \left| \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\theta(z, z_o)\hat{L}_z} \right| \bar{\boldsymbol{r}}_\perp \right\rangle \left\langle \bar{\boldsymbol{r}}_\perp \left| \hat{\underline{U}}_p(z, z_o) \right| \boldsymbol{r}_{\perp,o} \right\rangle. \end{aligned} \tag{2.115}$$

First, let us note that $\left\langle \boldsymbol{r}_\perp \left| \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\theta(z, z_o)\hat{L}_z} \right| \bar{\boldsymbol{r}}_\perp \right\rangle$, the matrix element of the operator for rotation around the z-axis through an angle $\theta(z, z_o)$, is given by (see Appendix E for details)

$$\left\langle \boldsymbol{r}_\perp \left| \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\theta(z, z_o)\hat{L}_z} \right| \bar{\boldsymbol{r}}_\perp \right\rangle = \delta^2\left(\boldsymbol{r}_\perp(\theta(z, z_o)) - \bar{\boldsymbol{r}}_\perp\right)\,, \tag{2.116}$$

where

$$
\begin{aligned}
\boldsymbol{r}_\perp(\theta(z,z_o)) &= (x(\theta(z,z_o)),\, y(\theta(z,z_o))) \\
\begin{pmatrix} x(\theta(z,z_o)) \\ y(\theta(z,z_o)) \end{pmatrix} &= \begin{pmatrix} \cos\theta(z,z_o) & -\sin\theta(z,z_o) \\ \sin\theta(z,z_o) & \cos\theta(z,z_o) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} .
\end{aligned}
\tag{2.117}
$$

Substituting the result of (2.116) in (2.115), we get

$$
G_p(\boldsymbol{r}_\perp, z; \boldsymbol{r}_{\perp,o}, z_o) = e^{\frac{i}{\hbar}p_0(z-z_o)} \underline{G}_p(\boldsymbol{r}_\perp(\theta(z,z_o)), z; \boldsymbol{r}_{\perp,o}, z_o) , \tag{2.118}
$$

where

$$
\underline{G}_p(\boldsymbol{r}_\perp(\theta(z,z_o)), z; \boldsymbol{r}_{\perp,o}, z_o) = \left\langle \boldsymbol{r}_\perp(\theta(z,z_o)) \left| \hat{\underline{U}}_p(z,z_o) \right| \boldsymbol{r}_{\perp,o} \right\rangle \tag{2.119}
$$

is the Green's function corresponding to the time-dependent-oscillator-like Hamiltonian $\hat{\underline{\mathbf{H}}}_{o,p}$.

In the classical case the paraxial transfer map could be neatly expressed in terms of the two linearly independent solutions to the classical equations of motion governed by the paraxial Hamiltonian. Likewise it is possible to derive the exact expressions for the paraxial transfer operator and the Green's function in this case too. This is possible due to a general result valid for any time-dependent quadratic Hamiltonian (*e.g.*, see [40]). The general result is equivalent to taking care of the infinite series in (2.72) completely by summing it exactly. In the present case, note that this is possible because of the Lie algebraic structure generated by the three operators $\{r_\perp^2,\ \hat{p}_\perp^2,\ \boldsymbol{r}_\perp \cdot \hat{\boldsymbol{p}}_\perp + \hat{\boldsymbol{p}}_\perp \cdot \boldsymbol{r}_\perp\}$. The details of the calculation of the transfer operator $\hat{\underline{U}}_p(z,z_o)$ and the Green's function $\underline{G}_p(\boldsymbol{r}_\perp(\theta(z,z_o)), z; \boldsymbol{r}_{\perp,o}, z_o)$, are given in Appendix F. The results are

$$
\begin{aligned}
& \hat{\underline{U}}_p(z,z_o) \\
& \quad = \exp\left\{ -\frac{i}{\hbar} \left(\frac{\varphi(z,z_o)}{\sin\varphi(z,z_o)} \right) \left[\frac{h_p(z,z_o)}{2p_0} \hat{p}_\perp^2 \right.\right. \\
& \qquad + \frac{1}{4} \left(g_p(z,z_o) - h_p'(z,z_o) \right) \left(\boldsymbol{r}_\perp \cdot \hat{\boldsymbol{p}}_\perp + \hat{\boldsymbol{p}}_\perp \cdot \boldsymbol{r}_\perp \right) \\
& \qquad \left.\left. - \frac{1}{2} p_0 g_p'(z,z_o) r_\perp^2 \right] \right\}
\end{aligned}
\tag{2.120}
$$

$$
\begin{aligned}
&\underline{G}_p\left(\boldsymbol{r}_\perp\left(\theta\left(z,z_o\right)\right),z;\boldsymbol{r}_{\perp,o},z_o\right)\\
&\quad=\frac{1}{\mathrm{i}\lambda_0 h_p\left(z,z_o\right)}\exp\left\{\frac{\mathrm{i}\pi}{\lambda_0 h_p\left(z,z_o\right)}\left[g_p\left(z,z_o\right)r_{\perp,o}^2\right.\right.\\
&\qquad\left.\left.-2\boldsymbol{r}_{\perp,o}\cdot\boldsymbol{r}_\perp\left(\theta\left(z,z_o\right)\right)+h_p'\left(z,z_o\right)r_\perp^2\right]\right\},\\
&\qquad\text{if}\ \ h_p\left(z,z_o\right)\neq 0
\end{aligned}
\tag{2.121}
$$

$$
\begin{aligned}
&\underline{G}_p\left(\boldsymbol{r}_\perp\left(\theta\left(z,z_o\right)\right),z;\boldsymbol{r}_{\perp,o},z_o\right)\\
&\quad=\frac{1}{g_p\left(z,z_o\right)}\exp\left\{\frac{\mathrm{i}\pi g_p'\left(z,z_o\right)}{\lambda_0 g_p\left(z,z_o\right)}r_\perp^2\right\}\\
&\qquad\times\delta^2\left(\boldsymbol{r}_{\perp,o}-\boldsymbol{r}_\perp\left(\theta\left(z,z_o\right)\right)/g_p\left(z,z_o\right)\right),\\
&\qquad\text{if}\ \ h_p\left(z,z_o\right)=0\,,
\end{aligned}
\tag{2.122}
$$

where

$$
\cos\varphi\left(z,z_o\right)=\frac{1}{2}\left(g_p\left(z,z_o\right)+h_p'\left(z,z_o\right)\right)\,,
\tag{2.123}
$$

with $g_p\left(z,z_o\right)$ and $h_p\left(z,z_o\right)$ as two linearly independent solutions of either (x or y) component of the equation

$$
\boldsymbol{r}_\perp''(z)+F(z)\boldsymbol{r}_\perp(z)=0\,,
\tag{2.124}
$$

satisfying the initial conditions

$$
g_p\left(z_o,z_o\right)=h_p'\left(z_o,z_o\right)=1\,,\quad h_p\left(z_o,z_o\right)=g_p'\left(z_o,z_o\right)=0\,,
\tag{2.125}
$$

and the relation

$$
g_p\left(z,z_o\right)h_p'\left(z,z_o\right)-h_p\left(z,z_o\right)g_p'\left(z,z_o\right)=1\,,\quad\text{for any}\ \ z\geq z_o\,.
\tag{2.126}
$$

Now, using the relations in (2.114)–(2.122), we get

$$
\begin{aligned}
&\psi(\boldsymbol{r}_\perp,z)\\
&\quad=\frac{e^{\mathrm{i}\gamma(z,z_o)}}{\mathrm{i}\lambda_0 h_p\left(z,z_o\right)}\int d^2r_o\,\exp\left\{\frac{\mathrm{i}\pi}{\lambda_0 h_p\left(z,z_o\right)}\left[g_p\left(z,z_o\right)r_{\perp,o}^2\right.\right.
\end{aligned}
$$

$$
\begin{aligned}
& \left. - 2\boldsymbol{r}_{\perp,\mathrm{o}} \cdot \boldsymbol{r}_{\perp} \left(\theta\left(z, z_{\mathrm{o}}\right)\right)\right]\Big\} \psi\left(\boldsymbol{r}_{\perp,\mathrm{o}}, z_{\mathrm{o}}\right) , \\
& \text{with } \gamma\left(z, z_{\mathrm{o}}\right) = \frac{2\pi}{\lambda_0}\left[\left(z - z_{\mathrm{o}}\right) + \frac{h_p'\left(z, z_{\mathrm{o}}\right)}{2h_p\left(z, z_{\mathrm{o}}\right)} r_{\perp}^2\right] , \\
& \text{if } h_p\left(z, z_{\mathrm{o}}\right) \neq 0
\end{aligned} \tag{2.127}
$$

$$
\begin{aligned}
\psi(\boldsymbol{r}_{\perp}, z) & = \frac{\mathrm{e}^{\mathrm{i}\gamma_0\left(z, z_{\mathrm{o}}\right)}}{g_p\left(z, z_{\mathrm{o}}\right)} \psi\left(\boldsymbol{r}_{\perp}\left(\theta\left(z, z_{\mathrm{o}}\right)\right) / g_p\left(z, z_{\mathrm{o}}\right), z_{\mathrm{o}}\right) , \\
& \text{with } \gamma_0\left(z, z_{\mathrm{o}}\right) = \frac{2\pi}{\lambda_0}\left[\left(z - z_{\mathrm{o}}\right) + \frac{g_p'\left(z, z_{\mathrm{o}}\right)}{2g_p\left(z, z_{\mathrm{o}}\right)} r_{\perp}^2\right] , \\
& \text{if } h_p\left(z, z_{\mathrm{o}}\right) = 0 ,
\end{aligned} \tag{2.128}
$$

representing the well-known general law of propagation of the paraxial beam wavefunction in the case of a round magnetic lens [4, 41, 42]. It may just be mentioned that equation (2.127) is the basis for the development of Fourier transform techniques in the theory of electron optical imaging process (for details, see [1]).

If $h_p\left(z, z_{\mathrm{o}}\right)$ vanishes at, say, $z = z_{\mathrm{i}}$, i.e., $h_p\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) = 0$, then, we can write

$$
\begin{aligned}
\psi\left(\boldsymbol{r}_{\perp,\mathrm{i}}, z_{\mathrm{i}}\right) & = \frac{1}{M} \mathrm{e}^{\mathrm{i}\gamma_0\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right)} \psi\left(\boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M, z_{\mathrm{o}}\right) , \\
& \text{with } M = g_p\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) , \quad \vartheta = \theta\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) \\
& \gamma_0\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) = \frac{2\pi}{\lambda_0}\left[\left(z_{\mathrm{i}} - z_{\mathrm{o}}\right) + g_p'\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) r_{\perp,\mathrm{i}}^2 / 2M\right]
\end{aligned} \tag{2.129}
$$

$$
\left|\psi\left(\boldsymbol{r}_{\perp,\mathrm{i}}, z_{\mathrm{i}}\right)\right|^2 = \frac{1}{M^2}\left|\psi\left(\boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M, z_{\mathrm{o}}\right)\right|^2 . \tag{2.130}
$$

This demonstrates that the plane at z_{i}, where $h_p\left(z, z_{\mathrm{o}}\right)$ vanishes, is the image plane and the intensity distribution at the object plane is reproduced exactly at the image plane with the magnification $M = g_p\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right)$ and the rotation through an angle

$$
\begin{aligned}
\vartheta & = \theta\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) = \int_{z_{\mathrm{o}}}^{z_{\mathrm{i}}} dz\, \theta'(z) \\
& = \frac{q}{2p_0} \int_{z_{\mathrm{o}}}^{z_{\mathrm{i}}} dz\, B(z) = \frac{q}{2p_0} \int_{z_{\mathrm{l}}}^{z_{\mathrm{r}}} dz\, B(z) .
\end{aligned} \tag{2.131}
$$

As is well-known, the general phenomenon of Larmor precession of any charged-particle in a magnetic field is responsible for this image rotation obtained in a single stage electron optical imaging using a round magnetic lens.

Since $\hat{\mathbf{H}}_{o,p}$ is hermitian the corresponding transfer operator will be unitary. Consequently, in this case, the total intensity of the beam at any plane is a conserved

$$
\begin{aligned}
\int d^2r_{\mathrm{i}}\, |\psi\left(\boldsymbol{r}_{\perp,\mathrm{i}}, z_{\mathrm{i}}\right)|^2 &= \frac{1}{M^2}\int d^2r_{\mathrm{i}}\, |\psi\left(\boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M, z_{\mathrm{o}}\right)|^2\,, \\
&= \frac{1}{M^2}\int d^2r_{\mathrm{i}}(\vartheta)\, |\psi\left(\boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M, z_{\mathrm{o}}\right)|^2\,, \\
&= \frac{1}{M^2}\int d^2r_{\mathrm{o}}\, |\psi\left(\boldsymbol{r}_{\perp,\mathrm{o}}/M, z_{\mathrm{o}}\right)|^2\,, \\
&= \int d^2r_{\mathrm{o}}\, |\psi\left(\boldsymbol{r}_{\perp,\mathrm{o}}, z_{\mathrm{o}}\right)|^2\,.
\end{aligned} \tag{2.132}
$$

We shall assume the strength of the lens field, or the value of $B(z)$, to be such that the first zero of $h_p\left(z, z_{\mathrm{o}}\right)$ is at $z = z_{\mathrm{i}} > z_r$. Then, as we shall see below, M is negative as should be in the case of a convergent lens forming a real inverted image.

So far, we have looked at imaging by paraxial beam from the point of view of the Schrödinger picture. Let us now look at this single stage Gaussian imaging using the Heisenberg picture, i.e., through the transfer maps $\left(\langle \boldsymbol{r}_\perp\rangle (z_{\mathrm{o}}), \langle \boldsymbol{p}_\perp\rangle (z_{\mathrm{o}})\right) \longrightarrow \left(\langle \boldsymbol{r}_\perp\rangle (z), \langle \boldsymbol{p}_\perp\rangle (z)\right)$. Using primarily (2.78) we get,

$$
\begin{aligned}
&\langle \boldsymbol{r}_\perp\rangle (z) \\
&= \left\langle \psi(z_{\mathrm{o}}) \left| \hat{U}_p^{\dagger}\left(z, z_{\mathrm{o}}\right) \boldsymbol{r}_\perp \hat{U}_p\left(z, z_{\mathrm{o}}\right) \right| \psi(z_{\mathrm{o}}) \right\rangle \\
&= \Big\langle \psi(z_{\mathrm{o}}) \Big| \underline{\hat{U}}_p^{\dagger}\left(z, z_{\mathrm{o}}\right) \mathrm{e}^{-\frac{i}{\hbar}\theta(z,z_{\mathrm{o}})\hat{L}_z} \boldsymbol{r}_\perp \\
&\qquad\qquad \times \mathrm{e}^{\frac{i}{\hbar}\theta(z,z_{\mathrm{o}})\hat{L}_z} \underline{\hat{U}}_p\left(z, z_{\mathrm{o}}\right) \Big| \psi(z_{\mathrm{o}}) \Big\rangle \\
&= \Big\langle \psi(z_{\mathrm{o}}) \Big| \mathrm{e}^{-\frac{i}{\hbar}\theta(z,z_{\mathrm{o}})\hat{L}_z} \underline{\hat{U}}_p^{\dagger}\left(z, z_{\mathrm{o}}\right) \boldsymbol{r}_\perp \\
&\qquad\qquad \times \underline{\hat{U}}_p\left(z, z_{\mathrm{o}}\right) \mathrm{e}^{\frac{i}{\hbar}\theta(z,z_{\mathrm{o}})\hat{L}_z} \Big| \psi(z_{\mathrm{o}}) \Big\rangle \\
&= \Big\langle \psi(z_{\mathrm{o}}) \Big| \mathrm{e}^{-\frac{i}{\hbar}\theta(z,z_{\mathrm{o}})\hat{L}_z} \left(g_p\left(z, z_{\mathrm{o}}\right) \boldsymbol{r}_\perp\right. \\
&\qquad\qquad \left. + h_p\left(z, z_{\mathrm{o}}\right) \hat{\boldsymbol{p}}_\perp / p_0\right) \mathrm{e}^{\frac{i}{\hbar}\theta(z,z_{\mathrm{o}})\hat{L}_z} \Big| \psi(z_{\mathrm{o}}) \Big\rangle \\
&= \langle \psi(z_{\mathrm{o}}) | g_p\left(z, z_{\mathrm{o}}\right) \boldsymbol{r}_\perp(-\theta\left(z, z_{\mathrm{o}}\right)) \\
&\qquad\qquad + h_p\left(z, z_{\mathrm{o}}\right) \hat{\boldsymbol{p}}_\perp(-\theta(z))/p_0 | \psi(z_{\mathrm{o}}) \rangle \\
&= g_p\left(z, z_{\mathrm{o}}\right) \langle \boldsymbol{r}_\perp(-\theta\left(z, z_{\mathrm{o}}\right)) \rangle (z_{\mathrm{o}})
\end{aligned}
$$

$$+h_p(z,z_o)\left\langle \boldsymbol{p}_\perp(-\theta(z,z_o))\right\rangle(z_o)/p_0\,, \qquad (2.133)$$

with

$$\begin{aligned}
&\left\langle \boldsymbol{r}_\perp(-\theta(z,z_o))\right\rangle(z_o)\\
&\quad = (\cos\theta(z,z_o)\langle x\rangle(z_o)+\sin\theta(z,z_o)\langle y\rangle(z_o)\,,\\
&\qquad -\sin\theta(z,z_o)\langle x\rangle(z_o)+\cos\theta(z,z_o)\langle y\rangle(z_o))\\
&\left\langle \boldsymbol{p}_\perp(-\theta(z,z_o))\right\rangle(z_o)\\
&\quad = (\cos\theta(z,z_o)\langle p_x\rangle(z_o)+\sin\theta(z,z_o)\langle p_y\rangle(z_o)\,,\\
&\qquad -\sin\theta(z,z_o)\langle p_x\rangle(z_o)+\cos\theta(z,z_o)\langle p_y\rangle(z_o))\,. \qquad (2.134)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\langle \boldsymbol{p}_\perp\rangle(z) &= p_0 g_p'(z,z_o)\left\langle \boldsymbol{r}_\perp(-\theta(z,z_o))\right\rangle(z_o)\\
&\quad +h_p'(z,z_o)\left\langle \boldsymbol{p}_\perp(-\theta(z,z_o))\right\rangle(z_o)\,. \qquad (2.135)
\end{aligned}$$

At the image plane at $z = z_i$, where $h_p(z_i, z_o) = 0$, the above transfer maps become

$$\begin{aligned}
\langle \boldsymbol{r}_\perp\rangle(z_i) &= M\left\langle \boldsymbol{r}_\perp(-\vartheta)\right\rangle(z_o)\\
\langle \boldsymbol{p}_\perp\rangle(z_i) &= p_0 g_p'(z_i,z_o)\left\langle \boldsymbol{r}_\perp(-\vartheta)\right\rangle(z_o)+\left\langle \boldsymbol{p}_\perp(-\vartheta)\right\rangle(z_o)/M\,, \qquad (2.136)
\end{aligned}$$

where ϑ is given in (2.131), $M = g_p(z_i, z_o)$ and $1/M = h_p'(z_i, z_o)$. The above equation implies that at the image plane a point-to-point, or stigmatic, image of the object is obtained and the image is magnified M times and rotated through an angle ϑ:

$$\begin{aligned}
\langle \boldsymbol{r}_\perp\rangle(z_i) &= \int d^2r_i\,\boldsymbol{r}_{\perp,i}\left|\psi(\boldsymbol{r}_{\perp,i},z_i)\right|^2\\
&= \frac{1}{M^2}\int d^2r_i\,\boldsymbol{r}_{\perp,i}\left|\psi(\boldsymbol{r}_{\perp,i}(\vartheta)/M,z_o)\right|^2\\
&= M\int d^2r_o\,\boldsymbol{r}_{\perp,o}(-\vartheta)\left|\psi(\boldsymbol{r}_{\perp,o},z_o)\right|^2\\
&= M\left\langle \boldsymbol{r}_\perp(-\vartheta)\right\rangle(z_o)\,. \qquad (2.137)
\end{aligned}$$

Let us now see how $\langle \boldsymbol{r}_\perp \rangle (z)$ and $\langle \boldsymbol{p}_\perp \rangle (z)$ evolve along the z-axis. Since

$$\begin{aligned}
\frac{\partial}{\partial z}\hat{U}_p(z,z_o) &= -\frac{\mathrm{i}}{\hbar}\hat{\mathbf{H}}_o\hat{U}_p(z,z_o) \\
\frac{\partial}{\partial z}\hat{U}_p^\dagger(z,z_o) &= \frac{\mathrm{i}}{\hbar}\hat{U}_p^\dagger(z,z_o)\,\hat{\mathbf{H}}_o\,,
\end{aligned} \tag{2.138}$$

it follows that

$$\frac{d}{dz}\langle \boldsymbol{r}_\perp\rangle(z) = \frac{\mathrm{i}}{\hbar}\left\langle \psi(z_o)\left|\hat{U}_p^\dagger(z,z_o)\left[\hat{\mathbf{H}}_{o,p},\boldsymbol{r}_\perp\right]\hat{U}_p(z,z_o)\right|\psi(z_o)\right\rangle \tag{2.139}$$

$$\frac{d}{dz}\langle \boldsymbol{p}_\perp\rangle(z) = \frac{\mathrm{i}}{\hbar}\left\langle \psi(z_o)\left|\hat{U}_p^\dagger(z,z_o)\left[\hat{\mathbf{H}}_{o,p},\hat{\boldsymbol{p}}_\perp\right]\hat{U}_p(z,z_o)\right|\psi(z_o)\right\rangle . \tag{2.140}$$

Explicitly, these equations of motion (2.139) and (2.140), become

$$\begin{aligned}
\frac{d}{dz}\begin{pmatrix}\langle x\rangle(z)\\ \langle y\rangle(z)\\ \langle p_x\rangle(z)/p_0\\ \langle p_y\rangle(z)/p_0\end{pmatrix} &= \tau(z)\begin{pmatrix}\langle x\rangle(z)\\ \langle y\rangle(z)\\ \langle p_x\rangle(z)/p_0\\ \langle p_y\rangle(z)/p_0\end{pmatrix} \\
\tau(z) &= [\underline{\tau}(z)+\rho(z)]
\end{aligned} \tag{2.141}$$

$$\underline{\tau}(z) = \begin{pmatrix} 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -F(z) & 0 & 0 & 0\\ 0 & -F(z) & 0 & 0\end{pmatrix} \tag{2.142}$$

$$\rho(z) = \begin{pmatrix} 0 & \theta'(z) & 0 & 0\\ -\theta'(z) & 0 & 0 & 0\\ 0 & 0 & 0 & \theta'(z)\\ 0 & 0 & -\theta'(z) & 0\end{pmatrix} . \tag{2.143}$$

Note that $\underline{\tau}(z)$ and $\rho(z)$ commute with each other as a result of the rotational symmetry of the system about the z-axis. As a consequence, we can write the solution for (2.141) as

$$\begin{aligned}
\begin{pmatrix}\langle x\rangle(z)\\ \langle y\rangle(z)\\ \langle p_x\rangle(z)/p_0\\ \langle p_y\rangle(z)/p_0\end{pmatrix} &= \mathcal{T}_p(z,z_o)\begin{pmatrix}\langle x\rangle(z_o)\\ \langle y\rangle(z_o)\\ \langle p_x\rangle(z_o)/p_0\\ \langle p_y\rangle(z_o)/p_0\end{pmatrix} \\
\mathcal{T}_p(z,z_o) &= \underline{\mathcal{T}}_p(z,z_o)\mathcal{R}(z,z_o) = \mathcal{R}(z,z_o)\underline{\mathcal{T}}_p(z,z_o)
\end{aligned} \tag{2.144}$$

where

$$\frac{d}{dz}\mathcal{R}(z,z_o) = \rho(z)\mathcal{R}(z,z_o)\,,\quad \mathcal{R}(z_o,z_o) = I \tag{2.145}$$

$$\frac{d}{dz}\mathcal{I}_p(z,z_o) = \underline{\tau}(z)\mathcal{I}_p(z,z_o)\,, \quad \mathcal{I}_p(z_o,z_o) = I\,, \tag{2.146}$$

and I is the 4×4 identity matrix.

Equation (2.145) can be readily integrated to give

$$\mathcal{R}(z,z_o) = \begin{pmatrix} \cos\theta\,(z,z_o) & \sin\theta\,(z,z_o) & 0 & 0 \\ -\sin\theta\,(z,z_o) & \cos\theta\,(z,z_o) & 0 & 0 \\ 0 & 0 & \cos\theta\,(z,z_o) & \sin\theta\,(z,z_o) \\ 0 & 0 & -\sin\theta\,(z,z_o) & \cos\theta\,(z,z_o) \end{pmatrix}. \tag{2.147}$$

If we now go to a rotated coordinate system such that we can write

$$\begin{pmatrix} \langle x\rangle(z) \\ \langle y\rangle(z) \\ \langle p_x\rangle(z) \\ \langle p_y\rangle(z) \end{pmatrix} = \mathcal{R}(z,z_o)\begin{pmatrix} \langle X\rangle(z) \\ \langle Y\rangle(z) \\ \langle P_X\rangle(z) \\ \langle P_Y\rangle(z) \end{pmatrix}, \tag{2.148}$$

with (X,Y) and (P_X,P_Y) respectively as the components of position and momentum in the new coordinate frame, then equation (2.144) takes the form

$$\begin{pmatrix} \langle X\rangle(z) \\ \langle Y\rangle(z) \\ \langle P_X\rangle(z)/p_0 \\ \langle P_Y\rangle(z)/p_0 \end{pmatrix} = \mathcal{I}_p(z,z_o)\begin{pmatrix} \langle X\rangle(z_o) \\ \langle Y\rangle(z_o) \\ \langle P_X\rangle(z_o)/p_0 \\ \langle P_Y\rangle(z_o)/p_0 \end{pmatrix}. \tag{2.149}$$

Note that xyz and XYz frames coincide at the object plane $(z = z_o)$. Then, the equations of motion for $\langle \boldsymbol{R}_\perp\rangle(z) = (\langle X\rangle(z), \langle Y\rangle(z))$ and $\langle \boldsymbol{P}_\perp\rangle(z) = (\langle P_X\rangle(z), \langle P_Y\rangle(z))$ become

$$\frac{d}{dz}\begin{pmatrix} \langle \boldsymbol{R}_\perp\rangle(z) \\ \langle \boldsymbol{P}_\perp\rangle(z)/p_0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -F(z) & 0 \end{pmatrix}\begin{pmatrix} \langle \boldsymbol{R}_\perp\rangle(z) \\ \langle \boldsymbol{P}_\perp\rangle(z)/p_0 \end{pmatrix}. \tag{2.150}$$

From (2.150) it follows that

$$\frac{d^2}{dz^2}\begin{pmatrix} \langle \boldsymbol{R}_\perp\rangle(z) \\ \langle \boldsymbol{P}_\perp\rangle(z)/p_0 \end{pmatrix} = \begin{pmatrix} -F(z) & 0 \\ -F'(z) & -F(z) \end{pmatrix}\begin{pmatrix} \langle \boldsymbol{R}_\perp\rangle(z) \\ \langle \boldsymbol{P}_\perp\rangle(z)/p_0 \end{pmatrix}. \tag{2.151}$$

or

$$\langle \boldsymbol{R}_\perp\rangle''(z) + F(z)\,\langle \boldsymbol{R}_\perp\rangle(z) = 0 \tag{2.152}$$

$$\frac{1}{p_0}\langle \boldsymbol{P}_\perp\rangle''(z) + F'(z)\,\langle \boldsymbol{R}_\perp\rangle(z) + \frac{1}{p_0}F(z)\,\langle \boldsymbol{P}_\perp\rangle(z) = 0 \tag{2.153}$$

which represent the paraxial equations of motion with reference to the rotated coordinate frame; now, compare (2.152) with (2.124). Equation (2.153) is not independent of (2.152) since it is just the consequence of the relation

$$\frac{d}{dz}\left\langle \boldsymbol{R}_\perp \right\rangle(z) = \frac{1}{p_0}\left\langle \boldsymbol{P}_\perp \right\rangle(z) \tag{2.154}$$

(see (2.150)), and a solution for $\left\langle \boldsymbol{R}_\perp \right\rangle(z)$ yields a solution for $\left\langle \boldsymbol{P}_\perp \right\rangle(z)$.

Equation (2.150) suggests that, due to its linearity, we can write its solution, in general, as

$$\begin{pmatrix} \left\langle \boldsymbol{R}_\perp \right\rangle(z) \\ \left\langle \boldsymbol{P}_\perp \right\rangle(z)/p_0 \end{pmatrix} = \begin{pmatrix} g_p(z, z_o) & h_p(z, z_o) \\ g_p'(z, z_o) & h_p'(z, z_o) \end{pmatrix} \begin{pmatrix} \left\langle \boldsymbol{R}_\perp \right\rangle(z_o) \\ \left\langle \boldsymbol{P}_\perp \right\rangle(z_o)/p_0 \end{pmatrix}, \tag{2.155}$$

where, as already mentioned above, the second relation follows from the first assumption in view of the first relation of (2.150), namely, $\frac{d}{dz}\left\langle \boldsymbol{R}_\perp \right\rangle(z) = \left\langle \boldsymbol{P}_\perp \right\rangle(z)/p_0$. Substituting the first relation of (2.155) in (2.152) it follows from the independence of $\left\langle \boldsymbol{R}_\perp \right\rangle(z_o)$ and $\left\langle \boldsymbol{P}_\perp \right\rangle(z_o)$ that

$$g_p''(z, z_o) + F(z) g_p(z, z_o) = 0\,, \quad h_p''(z, z_o) + F(z) h_p(z, z_o) = 0\,. \tag{2.156}$$

Since at $z = z_o$ the matrix in (2.155) should become identity we get the initial conditions for $g_p(z, z_o)$ and $h_p(z, z_o)$ as

$$g_p(z_o, z_o) = h_p'(z_o, z_o) = 1\,, \quad h_p(z_o, z_o) = g_p'(z_o, z_o) = 0\,. \tag{2.157}$$

In other words, $g_p(z, z_o)$ and $h_p(z, z_o)$ are two linearly independent solutions of either (X or Y) component of (2.152) subject to the initial conditions in (2.157). From the constancy of the Wronskian of any pair of independent solutions of a second order differential equation of the type in (2.152) we get

$$\begin{aligned} & g_p(z, z_o)\, h_p'(z, z_o) - h_p(z, z_o)\, g_p'(z, z_o) \\ & \quad = g_p(z_o, z_o) h_p'(z_o, z_o) - h_p(z_o, z_o) g_p'(z_o, z_o) = 1\,, \\ & \qquad \text{for any} \quad z \geq z_o\,. \end{aligned} \tag{2.158}$$

36818

INSTITUTE OF MATHEMATICAL SCIENCES LIB.

Thus, it is seen that the solutions of (2.124), $g_p(z, z_o)$ and $h_p(z, z_o)$, contained in (2.120)–(2.122), (2.127) and (2.128) can be obtained by integrating (2.150). Note that we can formally integrate (2.150) by applying the formula in (2.67) in view of the analogy between (2.60) and (2.150): the matrix in (2.155) can be obtained using (2.67) by replacing $\left(-\frac{i}{\hbar}\hat{\mathcal{H}}_o\right)$ by the matrix in (2.150). The result obtained gives $g_p(z, z_o)$ and $h_p(z, z_o)$ as infinite series expressions in terms of $F(z)$. Then, with

$$\mathcal{I}_p(z, z_o) = \begin{pmatrix} g_p(z, z_o) & 0 & h_p(z, z_o) & 0 \\ 0 & g_p(z, z_o) & 0 & h_p(z, z_o) \\ g_p'(z, z_o) & 0 & h_p'(z, z_o) & 0 \\ 0 & g_p'(z, z_o) & 0 & h_p'(z, z_o) \end{pmatrix}, \qquad (2.159)$$

and $\mathcal{R}(z, z_o)$ as given by (2.147), (2.144) is seen to be the matrix form of (2.133)–(2.135). This establishes the correspondence between the transfer operators in the Schrödinger picture and the transfer matrices in the Heisenberg picture :

$$\begin{aligned} e^{\frac{i}{\hbar}\theta(z,z_o)\hat{L}_z} &\longrightarrow \mathcal{R}(z, z_o), \quad \hat{\underline{U}}_p(z, z_o) \longrightarrow \mathcal{I}_p(z, z_o) \\ \hat{U}_p(z, z_o) &= e^{\frac{i}{\hbar}\theta(z,z_o)\hat{L}_z}\hat{\underline{U}}_p(z, z_o) \longrightarrow \mathcal{T}_p(z, z_o) \\ &= \mathcal{R}(z, z_o)\mathcal{I}_p(z, z_o) = \mathcal{I}_p(z, z_o)\mathcal{R}(z, z_o). \end{aligned} \qquad (2.160)$$

Explicitly,

$$\begin{aligned} & g_p(z, z_o) \\ & \quad = 1 - \int_{z_o}^{z} dz_1 \int_{z_o}^{z_1} dz F(z) \\ & \qquad + \int_{z_o}^{z} dz_3 \int_{z_o}^{z_3} dz_2 F(z_2) \int_{z_o}^{z_2} dz_1 \int_{z_o}^{z_1} dz F(z) - \ldots \\ & \qquad + \left\{ (-1)^n \int_{z_o}^{z} dz_{2n-1} \int_{z_o}^{z_{2n-1}} dz_{2n-2} F(z_{2n-2}) \right. \\ & \qquad \times \int_{z_o}^{z_{2n-2}} dz_{2n-3} \int_{z_o}^{z_{2n-3}} dz_{2n-4} F(z_{2n-4}) \ldots \\ & \qquad \left. \times \int_{z_o}^{z_2} dz_1 \int_{z_o}^{z_1} dz F(z) \right\} + \ldots \end{aligned} \qquad (2.161)$$

$$\begin{aligned} & h_p(z, z_o) \\ & \quad = (z - z_o) - \int_{z_o}^{z} dz_2 \int_{z_o}^{z_2} dz_1 F(z_1)(z_1 - z_o) \end{aligned}$$

$$
\begin{aligned}
&+\int_{z_o}^{z} dz_4 \int_{z_o}^{z_4} dz_3 F\left(z_3\right) \\
&\quad\times \int_{z_o}^{z_3} dz_2 \int_{z_o}^{z_2} dz_1 F\left(z_1\right)\left(z_1-z_o\right)-\ldots \\
&+\left\{(-1)^n \int_{z_o}^{z} dz_{2n} \int_{z_o}^{z_{2n}} dz_{2n-1} F\left(z_{2n-1}\right)\right. \\
&\quad\times \int_{z_o}^{z_{2n-1}} dz_{2n-2} \int_{z_o}^{z_{2n-2}} dz_{2n-3} F\left(z_{2n-3}\right) \ldots \\
&\quad\left.\times \int_{z_o}^{z_3} dz_2 \int_{z_o}^{z_2} dz_1 F\left(z_1\right)\left(z_1-z_o\right)\right\}+\ldots
\end{aligned}
\tag{2.162}
$$

$$
\begin{aligned}
&g_p'\left(z, z_o\right) \\
&=-\int_{z_o}^{z} dz F(z)+\int_{z_o}^{z} dz_2 F\left(z_2\right) \int_{z_o}^{z_2} dz_1 \int_{z_o}^{z_1} dz F(z) \\
&\quad-\ldots+\left\{(-1)^{n+1} \int_{z_o}^{z} dz_{2n} F\left(z_{2n}\right)\right. \\
&\quad\times \int_{z_o}^{z_{2n}} dz_{2n-1} \int_{z_o}^{z_{2n-1}} dz_{2n-2} F\left(z_{2n-2}\right) \ldots \\
&\quad\left.\times \int_{z_o}^{z_2} dz_1 \int_{z_o}^{z_1} dz F(z)\right\}+\ldots
\end{aligned}
\tag{2.163}
$$

$$
\begin{aligned}
&h_p'\left(z, z_o\right) \\
&=1-\int_{z_o}^{z} dz_1 F\left(z_1\right)\left(z_1-z_o\right) \\
&\quad+\int_{z_o}^{z} dz_3 F\left(z_3\right) \int_{z_o}^{z_3} dz_2 \int_{z_o}^{z_2} dz_1 F\left(z_1\right)\left(z_1-z_o\right)-\ldots \\
&\quad+\left\{(-1)^n \int_{z_o}^{z} dz_{2n-1} F\left(z_{2n-1}\right)\right. \\
&\quad\times \int_{z_o}^{z_{2n-1}} dz_{2n-2} \int_{z_o}^{z_{2n-2}} dz_{2n-3} F\left(z_{2n-3}\right) \ldots \\
&\quad\left.\times \int_{z_o}^{z_3} dz_2 \int_{z_o}^{z_2} dz_1 F\left(z_1\right)\left(z_1-z_o\right)\right\}+\ldots .
\end{aligned}
\tag{2.164}
$$

It is easy to verify directly that these expressions for $g_p\left(z, z_o\right)$ and $h_p\left(z, z_o\right)$ satisfy the equations (2.156) and (2.157).

The transfer operator defined by (2.66)–(2.69) (or (2.70)–(2.72)) is an ordered product of the transfer operators for successive infinitesimal distances from the initial point to the final point, an expression of the Huygens principle in operator form. Hence it can be written as an ordered product of transfer operators for successive finite distances covering the entire distance from the initial point to the final point.

Thus, we can write

$$\hat{U}_p(z, z_o) = \hat{U}_{D,p}(z, z_r)\,\hat{U}_{L,p}(z_r, z_l)\,\hat{U}_{D,p}(z_l, z_o)\,, \qquad \text{with } z_o < z_l\,,\ z > z_r\,, \tag{2.165}$$

where D refers to the drift in the field-free space and L refers to the propagation through the lens field. Consequently, one has

$$\begin{aligned}
&\psi(\boldsymbol{r}_\perp, z) \\
&= \int d^2r_r \int d^2r_l \int d^2r_o G_{D,p}(\boldsymbol{r}_\perp, z; \boldsymbol{r}_{\perp,r}, z_r) \\
&\qquad \times G_{L,p}(\boldsymbol{r}_{\perp,r}, z_r; \boldsymbol{r}_{\perp,l}, z_l) \\
&\qquad \times G_{D,p}(\boldsymbol{r}_{\perp,l}, z_l; \boldsymbol{r}_{\perp,o}, z_o)\,\psi(\boldsymbol{r}_{\perp,o}, z_o)\,.
\end{aligned} \tag{2.166}$$

Using the direct product notation for matrices,

$$\begin{aligned}
A \otimes B &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix},
\end{aligned} \tag{2.167}$$

the correspondence in (2.160) becomes, with $z_o < z_l$, $z > z_r$,

$$\begin{aligned}
\hat{U}_p(z, z_o) &\longrightarrow \mathcal{I}_p(z, z_o)\mathcal{R}(\theta(z, z_o)) = \mathcal{I}_p(z, z_o)\mathcal{R}(\vartheta) \\
&= \begin{pmatrix} g_p(z, z_o) & h_p(z, z_o) \\ g_p'(z, z_o) & h_p'(z, z_o) \end{pmatrix} \otimes R(\vartheta)
\end{aligned} \tag{2.168}$$

where $R(\vartheta) = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix}$, $\vartheta = \theta(z, z_o) = \theta(z_l, z_r)$. Since $F(z) = 0$ outside the lens region, we have, from equations (2.161)–(2.164),

$$\begin{aligned}
\hat{U}_{D,p}(z_l, z_o) &\longrightarrow \mathcal{I}_p(z_l, z_o)\,\mathcal{R}(\theta(z_l, z_o)) = \begin{pmatrix} 1 & z_l - z_o \\ 0 & 1 \end{pmatrix} \otimes \mathbb{1} \\
\hat{U}_{D,p}(z, z_r) &\longrightarrow \mathcal{I}_p(z, z_r)\,\mathcal{R}(\theta(z, z_r)) = \begin{pmatrix} 1 & z - z_r \\ 0 & 1 \end{pmatrix} \otimes \mathbb{1}
\end{aligned} \tag{2.169}$$

with $\mathbb{1}$ as the 2×2 identity matrix. For the lens region

$$
\begin{aligned}
\breve{U}_{L,p}(z_r, z_l) &\longrightarrow \mathcal{I}_p(z_r, z_l)\,\mathcal{R}(\theta(z_r, z_l)) = \mathcal{I}_p(z_r, z_l)\,\mathcal{R}(\vartheta) \\
&= \begin{pmatrix} g_{p,L} & h_{p,L} \\ g'_{p,L} & h'_{p,L} \end{pmatrix} \otimes R(\vartheta)\,,
\end{aligned} \tag{2.170}
$$

with $g_{p,L} = g_p(z_r, z_l)$, $h_{p,L} = h_p(z_r, z_l)$, $g'_{p,L} = g'_p(z, z_l)\,|_{z=z_r}$, $h'_{p,L} = h'_p(z, z_l)\,|_{z=z_r}$.
Then, substituting equations (2.169)–(2.170) in (2.165) we get the identity

$$
\begin{aligned}
&\mathcal{T}(z, z_o) \\
&= \begin{pmatrix} g_p(z, z_o) & h_p(z, z_o) \\ g'_p(z, z_o) & h'_p(z, z_o) \end{pmatrix} \otimes R(\vartheta) \\
&= \begin{pmatrix} g_{p,L} + (z - z_r)\, g'_{p,L} & \begin{array}{l} (z_l - z_o)\, g_{p,L} \\ +(z - z_r)(z_l - z_o)\, g'_{p,L} \\ +h_{p,L} + (z - z_r)\, h'_{p,L} \end{array} \\ & \\ g'_{p,L} & (z_l - z_o)\, g'_{p,L} + h'_{p,L} \end{pmatrix} \otimes R(\vartheta)\,.
\end{aligned} \tag{2.171}
$$

If the image plane is at $z = z_i$, then, from the vanishing of $h_p(z_i, z_o)$, we get

$$
(z_l - z_o)\, g_{p,L} + (z_i - z_r)(z_l - z_o)\, g'_{p,L} + h_{p,L} + (z_i - z_r)\, h'_{p,L} = 0\,. \tag{2.172}
$$

If we now substitute

$$
z_l - z_o = u - \frac{h'_{p,L} - 1}{g'_{p,L}}\,, \qquad z_i - z_r = v - \frac{g_{p,L} - 1}{g'_{p,L}}\,, \tag{2.173}
$$

then equation (2.172) becomes the familiar lens equation

$$
\frac{1}{u} + \frac{1}{v} = \frac{1}{f}\,, \tag{2.174}
$$

with the focal length f given by

$$
f = -\frac{1}{g'_{p,L}}\,. \tag{2.175}
$$

Equation (2.173) shows that the principal planes from which the object distance (u) and the image distance (v) are to be measured in the case of a thick lens are situated at

$$
\begin{aligned}
z_{\mathrm{P}_o} &= z_l + \frac{h'_{p,L} - 1}{g'_{p,L}} = z_l + f\left(1 - h'_{p,L}\right) \\
z_{\mathrm{P}_i} &= z_r - \frac{g_{p,L} - 1}{g'_{p,L}} = z_r - f\left(1 - g_{p,L}\right) .
\end{aligned}
\tag{2.176}
$$

The explicit expression for the focal length is now obtained from (2.163) and (2.175):

$$
\begin{aligned}
\frac{1}{f} &= \int_{z_l}^{z_r} dz\, F(z) - \int_{z_l}^{z_r} dz_2\, F\left(z_2\right) \int_{z_l}^{z_2} dz_1 \int_{z_l}^{z_1} dz\, F(z) + \ldots \\
&= \frac{q^2}{4p_0^2} \int_{z_l}^{z_r} dz\, B^2(z) \\
&\quad - \frac{q^4}{16p_0^4} \int_{z_l}^{z_r} dz_2\, B^2\left(z_2\right) \int_{z_l}^{z_2} dz_1 \int_{z_l}^{z_1} dz\, B^2(z) + \ldots .
\end{aligned}
\tag{2.177}
$$

To understand the behaviour of the above expression (2.177) for the focal length, let us consider the idealized model in which $B(z) = B =$ constant in the lens region and 0 outside. Then $1/f = (qB/2p_0)\sin\left(qBw/2p_0\right)$ where $w = (z_r - z_l)$ is the width, or thickness, of the lens. This shows that the focal length is always nonnegative to start with and is then periodic with respect to the variation of the field strength. Thus, the round magnetic lens is convergent up to a certain strength of the field beyond which it belongs to the class of divergent lenses though this terminology is never used due to the fact that the divergent character is really the result of very strong convergence (see [1] for more details). In practice, the common round magnetic lenses used in electron microscopy are convergent.

The paraxial transfer matrix from the object plane to the image plane now takes the form

$$
\mathcal{T}_p\left(z_i, z_o\right) = \begin{pmatrix} M & 0 \\ -1/f & 1/M \end{pmatrix} \otimes R(\vartheta) , \qquad M = -v/u , \tag{2.178}
$$

as is seen by simplifying (2.171) for $z = z_i$ using equations (2.172)–(2.175). Note that in our notation both u and v are positive and M is negative indicating the inverted nature of the image as should be in the case of imaging by a convergent lens. Another observation is in order. When the object is moved to $-\infty$, i.e., $u \longrightarrow \infty$, v is just f. Hence, the focus is situated at

$$z_F = z_{P_i} + f = z_r + f g_{p,L}\,. \tag{2.179}$$

Now, with the object situated at any $z_o < z_l$ the transfer matrix from the object plane to the back focal plane becomes

$$\mathcal{T}_p\left(z_F, z_o\right) = \begin{pmatrix} 0 & f \\ -\frac{1}{f} & 1 - \frac{u}{f} \end{pmatrix} \otimes R(\vartheta)\,, \tag{2.180}$$

as is seen by substituting $z = z_F$ in (2.171) and simplifying using (2.173), (2.175) and (2.179). The corresponding wave transfer relation in (2.127) shows that, apart from unimportant phase factor and constant multiplicative factor, the wavefunction in the back focal plane is equal to an inverse Fourier transform of the object wavefunction at $z_o < z_l$ (see [5] for more details).

Let us now consider the lens field to be weak such that

$$\int_{z_l}^{z_r} dz\, F(z) \ll 1/w\,, \qquad w = z_r - z_l\,. \tag{2.181}$$

Note that $\int_{z_l}^{z_r} dz\, F(z)$ has the dimension of reciprocal length and for the weak lens it is considered to be very small compared to the reciprocal of the characteristic length of the lens, namely, its width. In such a case, the formula for the focal length (2.177) can be approximated to give

$$\frac{1}{f} \approx \int_{z_l}^{z_r} dz\, F(z) = \frac{q^2}{4p_0^2}\int_{z_l}^{z_r} dz\, B^2(z) = \frac{q^2}{4p_0^2}\int_{-\infty}^{\infty} dz\, B^2(z) \tag{2.182}$$

which, first derived by Busch [43, 44], is known as Busch's formula for a thin axially symmetric magnetic lens (see [1], Chapters 16-17 for details on the classical theory of paraxial, or Gaussian, optics of rotationally symmetric electron lenses). A weak lens is said to be thin since in this case

$$f \gg w \tag{2.183}$$

as seen from equations (2.181) and (2.182).

For the thin lens the transfer matrix can be approximated as

$$
\begin{aligned}
\mathcal{T}_p(z, z_o) &\approx \mathcal{T}_{p,TL}(z, z_o) \\
&= \begin{pmatrix} 1 - \frac{1}{f}(z - z_P) & \begin{array}{c}(z - z_o) \\ -\frac{1}{f}(z_P - z_o)(z - z_P)\end{array} \\ -\frac{1}{f} & 1 - \frac{1}{f}(z_P - z_o) \end{pmatrix} \otimes R(\vartheta) \\
&= \left(\begin{pmatrix} 1 & z - z_P \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & z_P - z_o \\ 0 & 1 \end{pmatrix} \right) \otimes R(\vartheta)
\end{aligned}
$$

$$
\text{with} \quad z_P = \frac{1}{2}(z_l + z_r) \,. \tag{2.184}
$$

In this case the two principal planes collapse into a single principal plane at the center of the lens. If imaging occurs at $z = z_i$ for a given z_o then $u = z_P - z_o$ and $v = z_i - z_P$ satisfy the lens equation $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$ and the transfer matrix from the object plane to the image plane becomes

$$
\begin{pmatrix} M & 0 \\ -1/f & 1/M \end{pmatrix} \otimes R(\vartheta) \,, \qquad \text{with } M = -v/u \,. \tag{2.185}
$$

From the structure of the transfer matrix in (2.184) it is clear that apart from rotation and drifts through field-free regions in the front and back of the lens the effect of a thin lens is essentially described by the transfer matrix

$$
\begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \tag{2.186}
$$

which, as seen from (2.128), corresponds to multiplication of the wavefunction by the phase factor $\exp\left(-\frac{i\pi}{\lambda_0 f} r_\perp^2\right)$ as is well-known.

So far we have discussed the application of our formalism to the paraxial case. No realistic system is free from deviations (aberrations) from the ideal paraxial behaviour. So, we examine the beam-optical Hamiltonian beyond the paraxial approximation. The beam-optical Hamiltonian in (2.62) is of order four in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$.

The procedure we adopted to obtain (2.62) can be used to derive the beam-optical Hamiltonian to any order of accuracy. Let $\hat{\mathcal{H}}_0$ be the beam-optical Hamiltonian to desired order of accuracy say, n. We write $\hat{\mathcal{H}}_o$ as,

$$
\begin{aligned}
\hat{\mathcal{H}}_o &= \hat{\mathbf{H}}_{o,p} + \hat{\mathcal{H}}_{o,a} && (2.187)\\
\hat{\mathcal{H}}_{o,a} &= \hat{\mathbf{H}}_{o,a} + \hat{\mathcal{H}}_o^{(\lambda_0)} && (2.188)\\
\hat{\mathbf{H}}_{o,a} &= \sum_{m=3}^{n} \hat{\mathbf{H}}_{o,(m)} && (2.189)
\end{aligned}
$$

where $\hat{\mathbf{H}}_{o,p}$ governs the paraxial behaviour. $\hat{\mathcal{H}}_{o,a}$ gives rise to deviations from the paraxial behaviour. $\hat{\mathcal{H}}_{o,a}$ is the hermitian Hamiltonian of order three or more. $\hat{\mathcal{H}}_o^{(\lambda_0)}$, in general, is a sum of hermitian and antihermitian expressions with explicit λ_0-dependent terms of all possible orders (up to n), including linear and paraxial. $\hat{\mathbf{H}}_{o,(m)}$ are homogeneous polynomials of degree m in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$. We shall be treating all nonparaxial terms as aberrations irrespective of their type. It is to be noted that in the geometrical optics limit ($\lambda_0 \longrightarrow 0$) $\hat{\mathcal{H}}_o^{(\lambda_0)}$ vanishes and $\hat{\mathbf{H}}_{o,(m)}$ reproduces the classical counterpart $\mathcal{H}_{(m)}$. As in the classical case we shall resort to the interaction picture for studying the aberrations.

We are seeking the solutions of

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial z}|\psi(z)\rangle &= \hat{\mathcal{H}}_o|\psi(z)\rangle\\
\hat{\mathcal{H}}_o &= \hat{\mathbf{H}}_{o,p} + \hat{\mathcal{H}}_{o,a}\,, && (2.190)
\end{aligned}
$$

in the form

$$
\begin{aligned}
|\psi(z)\rangle &= \hat{\mathcal{T}}(z, z_o)|\psi(z_o)\rangle\,,\\
&\qquad \text{where } \hat{\mathcal{T}}(z_o, z_o) = \hat{\mathcal{I}} && (2.191)
\end{aligned}
$$

with

$$
i\hbar\frac{\partial}{\partial z}\hat{\mathcal{T}}(z, z_o) = \hat{\mathcal{H}}_o\hat{\mathcal{T}}(z, z_o) \qquad (2.192)
$$

and

$$\mathrm{i}\hbar\frac{\partial}{\partial z}\hat{U}_p(z,z_o) = \hat{\mathbf{H}}_{o,p}\hat{U}_p(z,z_o) \tag{2.193}$$

Let

$$|\psi(z)\rangle = \hat{U}_p(z,z_o)\left|\psi^I(z)\right\rangle \tag{2.194}$$

so that equation (2.190) becomes

$$\begin{aligned}
\mathrm{i}\hbar\frac{\partial}{\partial z}|\psi(z)\rangle &= \mathrm{i}\hbar\frac{\partial}{\partial z}\left\{\hat{U}_p(z,z_o)\left|\psi^I(z)\right\rangle\right\} \\
&\left\{\left(\mathrm{i}\hbar\frac{\partial}{\partial z}\hat{U}_p(z,z_o)\right) + \hat{U}_p(z,z_o)\,\mathrm{i}\hbar\frac{\partial}{\partial z}\right\}|\psi^I(z)\rangle \\
&= \left\{\hat{\mathbf{H}}_{o,p}\hat{U}_p(z,z_o) + \hat{U}_p(z,z_o)\,\mathrm{i}\hbar\frac{\partial}{\partial z}\right\}|\psi^I(z)\rangle \\
&= \hat{\mathbf{H}}_{o,p}|\psi(z)\rangle + \hat{U}_p(z,z_o)\,\mathrm{i}\hbar\frac{\partial}{\partial z}|\psi^I(z)\rangle .
\end{aligned} \tag{2.195}$$

From (2.190) and (2.195), we get

$$\begin{aligned}
\mathrm{i}\hbar\frac{\partial}{\partial z}|\psi^I(z)\rangle &= \hat{U}_p^\dagger(z,z_o)\left(\hat{\mathcal{H}}_o - \hat{\mathbf{H}}_{o,p}\right)\hat{U}_p(z,z_o)\,\psi^I(z)\rangle \\
&= \hat{\mathcal{H}}_{o,a}^I|\psi^I(z)\rangle ,
\end{aligned} \tag{2.196}$$

with

$$\hat{\mathcal{H}}_{o,a}^I(z) = \hat{U}_p^\dagger(z,z_o)\,\hat{\mathcal{H}}_{o,a}(z)\,\hat{U}_p(z,z_o)\ , \tag{2.197}$$

where the superscript I denotes the so-called interaction picture.

Integrating (2.196), we have

$$|\psi^I(z)\rangle = \hat{\mathcal{T}}^I(z,z_o)|\psi^I(z_o)\rangle \tag{2.198}$$

with

$$\hat{\mathcal{T}}^I(z,z_o) = \wp\left\{\exp\left(-\frac{\mathrm{i}}{\hbar}\int_{z_o}^{z} dz\,\hat{\mathcal{H}}_{o,a}^I(z)\right)\right\}\ , \tag{2.199}$$

From equations (2.194) and (2.198) we have

$$\hat{\mathcal{T}}(z,z_o) = \hat{U}_p^\dagger(z,z_o)\,\hat{\mathcal{T}}^I(z,z_o)\hat{U}_p(z,z_o)\ , \tag{2.200}$$

generalizing the paraxial law by including the aberrations. It is to be noted that $\hat{\mathcal{T}}^I(z, z_o)$ has to be evaluated to an order consistent with the evaluation of the beam-optical Hamiltonian. To the desired order, the transfer maps become

$$\begin{aligned} \langle \boldsymbol{r}_\perp \rangle (z) &= \left\langle \hat{\mathcal{T}}^{I\dagger} \hat{U}_p^\dagger \boldsymbol{r}_\perp \hat{U}_p \hat{\mathcal{T}}^I \right\rangle (z_o) \qquad (2.201) \\ \langle \boldsymbol{p}_\perp \rangle (z) &= \left\langle \hat{\mathcal{T}}^{I\dagger} \hat{U}_p^\dagger \hat{\boldsymbol{p}}_\perp \hat{U}_p \hat{\mathcal{T}}^I \right\rangle (z_o) . \qquad (2.202) \end{aligned}$$

Let us first consider the case of the magnetic round lens. We will approximate the beam-optical Hamiltonian by dropping the λ_0-dependent terms, and retain only the leading order contributions to aberrations. Thus,

$$\hat{\mathcal{H}}_o \approx \hat{\mathbf{H}}_o^{(4)} = \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,(4)} , \qquad (2.203)$$

and the corresponding transfer operator $\hat{\mathcal{T}}(z, z_o)$ is now unitary

$$\begin{aligned} \hat{U}^I_{(4)}(z, z_o) &= \wp \left\{ \exp \left(-\frac{\mathrm{i}}{\hbar} \int_{z_o}^{z} dz\, \hat{\mathbf{H}}^I_{o,(4)}(z) \right) \right\} , \\ &= \exp \left\{ -\frac{\mathrm{i}}{\hbar} \int_{z_o}^{z} dz\, \hat{\mathbf{H}}^I_{o,(4)}(z) \right\} , \qquad \text{to order four} \qquad (2.204) \end{aligned}$$

where we have disregarded all the commutator terms in the formula for $\hat{U}$ since they lead to polynomials of degree higher than four in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$.

Using the result

$$\begin{pmatrix} \hat{U}_p^\dagger x \hat{U}_p \\ \hat{U}_p^\dagger y \hat{U}_p \\ \hat{U}_p^\dagger (\hat{p}_x/p_0) \hat{U}_p \\ \hat{U}_p^\dagger (\hat{p}_y/p_0) \hat{U}_p \end{pmatrix} = \begin{pmatrix} g & h \\ g' & h' \end{pmatrix} \otimes R(\theta) \begin{pmatrix} x \\ y \\ \hat{p}_x/p_0 \\ \hat{p}_y/p_0 \end{pmatrix} \qquad (2.205)$$

(see (2.133)–(2.135)), with $\hat{U}_p = \hat{U}_p(z, z_o)$, $g = g_p(z, z_o)$, $h = h_p(z, z_o)$, $g' = g_p'(z, z_o)$, $h' = h_p'(z, z_o)$ and $\theta = \theta(z, z_o)$, and (2.103) and (2.204), we find, after considerable but straightforward algebra, that

$$\begin{aligned} \hat{U}^I_{(4)}(z, z_o) &= \exp \left\{ -\frac{\mathrm{i}}{\hbar} \hat{\mathbb{T}}^I_{(4)}(z, z_o) \right\} \\ &= \exp \left\{ -\frac{\mathrm{i}}{\hbar} \left[\frac{1}{4p_0^3} C(z, z_o) \hat{p}_\perp^4 \right. \right. \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{4p_0^2}K(z,z_o)\left\{\hat{p}_\perp^2, \hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\\
&+\frac{1}{p_0^2}k(z,z_o)\hat{p}_\perp^2\hat{L}_z\\
&+\frac{1}{4p_0}A(z,z_o)\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)^2\\
&+\frac{1}{2p_0}a(z,z_o)\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)\hat{L}_z\\
&+\frac{1}{4p_0}F(z,z_o)\left(\hat{p}_\perp^2 r_\perp^2+r_\perp^2\hat{p}_\perp^2\right)\\
&+\frac{1}{4}D(z,z_o)\left\{\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp, r_\perp^2\right\}\\
&+d(z,z_o)r_\perp^2\hat{L}_z\\
&+\frac{p_0}{4}E(z,z_o)r_\perp^4\Big]\Big\}\,, \qquad (2.206)
\end{aligned}
$$

where $\{A\,,\,B\} = AB + BA$ and

$$
\begin{aligned}
C(z,z_o) &= \frac{1}{2}\int_{z_o}^{z} dz\,\left\{\left(\alpha^4-\alpha\alpha''\right)h^4+2\alpha^2h^2h'^2+h'^4\right\}\\
K(z,z_o) &= \frac{1}{2}\int_{z_o}^{z} dz\,\left\{\left(\alpha^4-\alpha\alpha''\right)gh^3+\alpha^2(gh)'hh'+g'h'^3\right\}\\
k(z,z_o) &= \int_{z_o}^{z} dz\,\left\{\left(\frac{1}{8}\alpha''-\frac{1}{2}\alpha^3\right)h^2-\frac{1}{2}\alpha h'^2\right\}\\
A(z,z_o) &= \frac{1}{2}\int_{z_o}^{z} dz\,\left\{\left(\alpha^4-\alpha\alpha''\right)g^2h^2+2\alpha^2gg'hh'+g'^2h'^2-\alpha^2\right\}\\
a(z,z_o) &= \int_{z_o}^{z} dz\,\left\{\left(\frac{1}{4}\alpha''-\alpha^3\right)gh-\alpha g'h'\right\}\\
F(z,z_o) &= \frac{1}{2}\int_{z_o}^{z} dz\,\left\{\left(\alpha^4-\alpha\alpha''\right)g^2h^2+\alpha^2\left(g^2h'^2+g'^2h^2\right)+g'^2h'^2+2\alpha^2\right\}\\
D(z,z_o) &= \frac{1}{2}\int_{z_o}^{z} dz\,\left\{\left(\alpha^4-\alpha\alpha''\right)g^3h+\alpha^2gg'(gh)'+g'^3h'\right\}\\
d(z,z_o) &= \int_{z_o}^{z} dz\,\left\{\left(\frac{1}{8}\alpha''-\frac{1}{2}\alpha^3\right)g^2-\frac{1}{2}\alpha g'^2\right\}\\
E(z,z_o) &= \frac{1}{2}\int_{z_o}^{z} dz\,\left\{\left(\alpha^4-\alpha\alpha''\right)g^4+2\alpha g^2g'^2+g'^4\right\}\,. \qquad (2.207)
\end{aligned}
$$

From (2.194) and (2.198), we have

$$
|\psi(z)\rangle = \hat{U}_p\left(z,z_o\right)\hat{U}_{(4)}^{I}\left(z,z_o\right)|\psi(z_o)\rangle \qquad (2.208)
$$

which represents the generalization of the paraxial propagation law in (2.127), corresponding to the inclusion of the lowest order aberrations. Now, the transfer map becomes

$$\langle \boldsymbol{r}_\perp \rangle_{(3)}(z) = \left\langle \hat{U}^{I\dagger}_{(4)} \hat{U}^\dagger_p \boldsymbol{r}_\perp \hat{U}_p \hat{U}^I_{(4)} \right\rangle (z_o) \tag{2.209}$$

$$\langle \boldsymbol{p}_\perp \rangle_{(3)}(z) = \left\langle \hat{U}^{I\dagger}_{(4)} \hat{U}^\dagger_p \hat{\boldsymbol{p}}_\perp \hat{U}_p \hat{U}^I_{(4)} \right\rangle (z_o)\,, \tag{2.210}$$

with $\hat{U}^I_{(4)} = \hat{U}^I_{(4)}(z, z_o)$ and $\langle \ldots \rangle (z_o) = \langle \psi(z_o) | \ldots | \psi(z_o) \rangle$. The subscript (3) indicates that the correction to the paraxial (or, first-order) result incorporated involves up to third order polynomials in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$. Explicitly,

$$\begin{aligned}
\begin{pmatrix} \langle x \rangle_{(3)}(z) \\ \langle y \rangle_{(3)}(z) \\ \langle p_x \rangle_{(3)}(z)/p_0 \\ \langle p_y \rangle_{(3)}(z)/p_0 \end{pmatrix} &= \left(\begin{pmatrix} g & h \\ g' & h' \end{pmatrix} \otimes R(\theta) \right) \\
&\qquad \times \begin{pmatrix} \left\langle \hat{U}^{I\dagger}_{(4)} x \hat{U}^I_{(4)} \right\rangle (z_o) \\ \left\langle \hat{U}^{I\dagger}_{(4)} y \hat{U}^I_{(4)} \right\rangle (z_o) \\ \left\langle \hat{U}^{I\dagger}_{(4)} \hat{p}_x \hat{U}^I_{(4)} \right\rangle (z_o)/p_0 \\ \left\langle \hat{U}^{I\dagger}_{(4)} \hat{p}_y \hat{U}^I_{(4)} \right\rangle (z_o)/p_0 \end{pmatrix} \\
&= \begin{pmatrix} \langle x \rangle_p(z) \\ \langle y \rangle_p(z) \\ \langle p_x \rangle_p(z)/p_0 \\ \langle p_y \rangle_p(z)/p_0 \end{pmatrix} + \begin{pmatrix} (\Delta x)_{(3)}(z) \\ (\Delta y)_{(3)}(z) \\ (\Delta p_x)_{(3)}(z)/p_0 \\ (\Delta p_y)_{(3)}(z)/p_0 \end{pmatrix}
\end{aligned} \tag{2.211}$$

where the geometrical aberrations, or the deviations from the paraxial results, are given by

$$\begin{aligned}
\begin{pmatrix} (\Delta x)_{(3)}(z) \\ (\Delta y)_{(3)}(z) \\ (\Delta p_x)_{(3)}(z)/p_0 \\ (\Delta p_y)_{(3)}(z)/p_0 \end{pmatrix} &\approx \left(\begin{pmatrix} g & h \\ g' & h' \end{pmatrix} \otimes R(\theta) \right) \\
&\qquad \times \begin{pmatrix} \left\langle \frac{i}{\hbar} \left[\hat{\mathbb{T}}^I_{(4)}\,,\, x \right] \right\rangle (z_o) \\ \left\langle \frac{i}{\hbar} \left[\hat{\mathbb{T}}^I_{(4)}\,,\, y \right] \right\rangle (z_o) \\ \left\langle \frac{i}{\hbar} \left[\hat{\mathbb{T}}^I_{(4)}\,,\, \hat{p}_x \right] \right\rangle (z_o)/p_0 \\ \left\langle \frac{i}{\hbar} \left[\hat{\mathbb{T}}^I_{(4)}\,,\, \hat{p}_y \right] \right\rangle (z_o)/p_0 \end{pmatrix},
\end{aligned}$$

$$(2.212)$$

involving expectation values of homogeneous third order polynomials in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$. Hence the subscript (3) for $(\Delta x)_{(3)}(z)$, $(\Delta y)_{(3)}(z)$, etc., and the name third order aberrations. Note that, here, we are retaining only the single commutator terms in the application of the formula in (1.78) to compute $\hat{U}^{I\,\dagger}_{(4)} x \hat{U}^{I}_{(4)}$, $\hat{U}^{I\,\dagger}_{(4)} y \hat{U}^{I}_{(4)}$, etc., since the remaining multiple commutator terms lead to polynomials in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$ which are only of degree ≥ 5 and are to be ignored in order to be consistent with the fact that we have retained only terms up to fourth order in $(\boldsymbol{r}_\perp, \hat{\boldsymbol{p}}_\perp)$ in the Hamiltonian and the transfer operator $\hat{U}$.

Obviously, the plane at which the influence of aberrations is to be known is the image plane at $z = z_i$:

$$\begin{pmatrix} \langle x\rangle_{(3)}(z_i) \\ \langle y\rangle_{(3)}(z_i) \\ \langle p_x\rangle_{(3)}(z_i)/p_0 \\ \langle p_y\rangle_{(3)}(z_i)/p_0 \end{pmatrix} \approx \left(\begin{pmatrix} M & 0 \\ -1/f & 1/M \end{pmatrix} \otimes R(\vartheta) \right) \times \left(\begin{pmatrix} \langle x\rangle(z_o) \\ \langle y\rangle(z_o) \\ \langle p_x\rangle(z_o)/p_0 \\ \langle p_y\rangle(z_o)/p_0 \end{pmatrix} + \begin{pmatrix} (\delta x)_{(3)}(z_o) \\ (\delta y)_{(3)}(z_o) \\ (\delta p_x)_{(3)}(z_o)/p_0 \\ (\delta p_y)_{(3)}(z_o)/p_0 \end{pmatrix} \right) \qquad (2.213)$$

where

$$\begin{aligned} (\delta x)_{(3)}(z_o) = \; & C_s \left\langle \hat{p}_x \hat{p}_\perp^2 \right\rangle (z_o)/p_0^3 \\ & + K \left\langle \frac{1}{2} \left\{ \hat{p}_x \, , \, \hat{\boldsymbol{p}}_\perp \cdot \boldsymbol{r}_\perp + \boldsymbol{r}_\perp \cdot \hat{\boldsymbol{p}}_\perp \right\} \right. \\ & \qquad \left. + \frac{1}{2} \left\{ x \, , \, \hat{p}_\perp^2 \right\} \right\rangle (z_o)/p_0^2 \\ & + k \left\langle \left\{ \hat{p}_x \, , \, \hat{L}_z \right\} - \frac{1}{2} \left\{ y \, , \, \hat{p}_\perp^2 \right\} \right\rangle (z_o)/p_0^2 \\ & + A \left\langle \frac{1}{2} \left\{ x \, , \, \hat{\boldsymbol{p}}_\perp \cdot \boldsymbol{r}_\perp + \boldsymbol{r}_\perp \cdot \hat{\boldsymbol{p}}_\perp \right\} \right\rangle (z_o)/p_0 \\ & + a \left\langle \frac{1}{2} \left\{ x \, , \, \hat{L}_z \right\} \right. \\ & \qquad \left. - \frac{1}{2} \left\{ y \, , \, \hat{\boldsymbol{p}}_\perp \cdot \boldsymbol{r}_\perp + \boldsymbol{r}_\perp \cdot \hat{\boldsymbol{p}}_\perp \right\} \right\rangle (z_o)/p_0 \end{aligned}$$

$$
\begin{aligned}
&+F\left\langle\frac{1}{2}\left\{\hat{p}_x\,,\,r_\perp^2\right\}\right\rangle(z_o)/p_0 \\
&+D\left\langle x r_\perp^2\right\rangle(z_o) \\
&-d\left\langle y r_\perp^2\right\rangle(z_o)
\end{aligned}
\tag{2.214}
$$

$$
\begin{aligned}
(\delta y)_{(3)}(z_o) \;=\;& C_s\left\langle\hat{p}_y\hat{p}_\perp^2\right\rangle(z_o)/p_0^3 \\
&+K\left\langle\frac{1}{2}\left\{\hat{p}_y\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right. \\
&\qquad\left.+\frac{1}{2}\left\{y\,,\,\hat{p}_\perp^2\right\}\right\rangle(z_o)/p_0^2 \\
&+k\left\langle\left\{\hat{p}_y\,,\,\hat{L}_z\right\}+\frac{1}{2}\left\{x\,,\,\hat{p}_\perp^2\right\}\right\rangle(z_o)/p_0^2 \\
&+A\left\langle\frac{1}{2}\left\{y\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o)/p_0 \\
&+a\left\langle\frac{1}{2}\left\{y\,,\,\hat{L}_z\right\}\right. \\
&\qquad\left.-\frac{1}{2}\left\{x\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o)/p_0 \\
&+F\left\langle\frac{1}{2}\left\{\hat{p}_y\,,\,r_\perp^2\right\}\right\rangle(z_o)/p_0 \\
&+D\left\langle y r_\perp^2\right\rangle(z_o) \\
&+d\left\langle x r_\perp^2\right\rangle(z_o)
\end{aligned}
\tag{2.215}
$$

$$
\begin{aligned}
(\delta p_x)_{(3)}(z_o) \;=\;& -K\left\langle\hat{p}_x\hat{p}_\perp^2\right\rangle(z_o)/p_0^2 \\
&-k\left\langle\hat{p}_y\hat{p}_\perp^2\right\rangle(z_o)/p_0^2 \\
&-A\left\langle\frac{1}{2}\left\{\hat{p}_x\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o)/p_0 \\
&-a\left\langle\frac{1}{2}\left\{\hat{p}_x\,,\,\hat{L}_z\right\}\right. \\
&\qquad\left.+\frac{1}{2}\left\{\hat{p}_y\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o)/p_0 \\
&-F\left\langle\frac{1}{2}\left\{x\,,\,\hat{p}_\perp^2\right\}\right\rangle(z_o)/p_0 \\
&-D\left\langle\frac{1}{2}\left\{\hat{p}_x\,,\,r_\perp^2\right\}\right. \\
&\qquad\left.+\frac{1}{2}\left\{x\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o) \\
&-d\left\langle\left\{x\,,\,\hat{L}_z\right\}+\frac{1}{2}\left\{\hat{p}_y\,,\,r_\perp^2\right\}\right\rangle(z_o)
\end{aligned}
$$

$$-Ep_0\left\langle xr_\perp^2\right\rangle(z_o) \tag{2.216}$$

$$\begin{aligned}
(\delta p_y)_{(3)}(z_o) = & -K\left\langle \hat{p}_y\hat{p}_\perp^2\right\rangle(z_o)/p_0^2 \\
& +k\left\langle \hat{p}_x\hat{p}_\perp^2\right\rangle(z_o)/p_0^2 \\
& -A\left\langle\frac{1}{2}\left\{\hat{p}_y\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o)/p_0 \\
& -a\left\langle\frac{1}{2}\left\{\hat{p}_y\,,\,\hat{L}_z\right\}\right. \\
& \qquad \left.-\frac{1}{2}\left\{\hat{p}_x\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o)/p_0 \\
& -F\left\langle\frac{1}{2}\left\{y\,,\,\hat{p}_\perp^2\right\}\right\rangle(z_o)/p_0 \\
& -D\left\langle\frac{1}{2}\left\{\hat{p}_y\,,\,r_\perp^2\right\}\right. \\
& \qquad \left.+\frac{1}{2}\left\{y\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right\}\right\rangle(z_o) \\
& -d\left\langle\left\{y\,,\,\hat{L}_z\right\}-\frac{1}{2}\left\{\hat{p}_x\,,\,r_\perp^2\right\}\right\rangle(z_o) \\
& -Ep_0\left\langle yr_\perp^2\right\rangle(z_o)
\end{aligned} \tag{2.217}$$

with

$$\begin{aligned}
C_s &= C(z_i, z_o)\,, \quad K = K(z_i, z_o)\,, \quad k = k(z_i, z_o)\,, \\
A &= A(z_i, z_o)\,, \quad a = a(z_i, z_o)\,, \quad F = F(z_i, z_o)\,, \\
D &= D(z_i, z_o)\,, \quad d = d(z_i, z_o)\,, \quad E = E(z_i, z_o)\,.
\end{aligned} \tag{2.218}$$

With reference to the aberrations of position (see (2.214) and (2.215)) constants C_s, K, k, A, a, F, D and d are known as the aberration coefficients corresponding, respectively, to spherical aberration, coma, anisotropic coma, astigmatism, anisotropic astigmatism, curvature of field, distortion and anisotropic distortion (see [1], for a detailed picture of the effects of these geometrical aberrations on the quality of the image and the classical methods of computation of these aberrations; see Ximen [36, 37] for a treatment of the classical theory of geometrical aberrations using position, momentum and the Hamiltonian equations of motion). The gradient aberrations (see (2.216) and (2.217)) do not affect the single-stage image but should be taken into account as the input to the next stage when the lens forms a part of a complex imaging system.

It is interesting to note the following symmetry of the nine aberration coefficients : under the exchange $g \longleftrightarrow h$, the coefficients transform as $C_s \longleftrightarrow E$, $K \longleftrightarrow D$, $k \longleftrightarrow d$, $A \longleftrightarrow F$, and a remains invariant. To see the connection $A \longleftrightarrow F$ we have to use the relation $gh' - hg' = 1$.

Introducing the notations

$$u = x + \mathrm{i}y\,, \qquad v = \left(\hat{p}_x + \mathrm{i}\hat{p}_y\right)/p_0\,, \tag{2.219}$$

the above transfer maps can be written in a compact matrix form (see [1], Chapter 27, for the aberration matrices in the classical context) as follows:

$$\begin{aligned}
\begin{pmatrix} \langle u\rangle_{(3)}(z_{\mathrm{i}}) \\ \langle v\rangle_{(3)}(z_{\mathrm{i}}) \end{pmatrix} &= \mathrm{e}^{-\mathrm{i}\theta}\begin{pmatrix} M & 0 \\ -1/f & 1/M \end{pmatrix} \\
&\quad \times \begin{pmatrix} 1 & 0 & C_s & 2K & 2k & F & D+\mathrm{i}d & 2A+\mathrm{i}a & -a & K+\mathrm{i}k \\ 0 & 1 & \mathrm{i}k-K & \mathrm{i}a-2A & -a & \mathrm{i}d-D & -E & -2D & 2d & -F \end{pmatrix} \\
&\quad \times \begin{pmatrix} \langle u\rangle(z_{\mathrm{o}}) \\ \langle v\rangle(z_{\mathrm{o}}) \\ \langle vv^\dagger v\rangle(z_{\mathrm{o}}) \\ \frac{1}{4}\langle\left\{v\,,\, u^\dagger v + v^\dagger u\right\}\rangle(z_{\mathrm{o}}) \\ \frac{1}{4\mathrm{i}}\langle\left\{v\,,\, u^\dagger v - v^\dagger u\right\}\rangle(z_{\mathrm{o}}) \\ \frac{1}{2}\langle\left\{v\,,\, u^\dagger u\right\}\rangle(z_{\mathrm{o}}) \\ \langle uu^\dagger u\rangle(z_{\mathrm{o}}) \\ \frac{1}{4}\langle\left\{u\,,\, v^\dagger u + u^\dagger v\right\}\rangle(z_{\mathrm{o}}) \\ \frac{1}{4\mathrm{i}}\langle\left\{u\,,\, v^\dagger u - u^\dagger v\right\}\rangle(z_{\mathrm{o}}) \\ \frac{1}{2}\langle\left\{u\,,\, v^\dagger v\right\}\rangle(z_{\mathrm{o}}) \end{pmatrix} .
\end{aligned} \tag{2.220}$$

Let us now look at the wavefunction in the image plane. We have

$$\begin{aligned}
\psi(\boldsymbol{r}_{\perp,\mathrm{i}}, z_{\mathrm{i}}) &= \int d^2 r_{\mathrm{o}} \int d^2 r^{(\mathrm{i})} \left\langle \boldsymbol{r}_{\perp,\mathrm{i}} \left| \hat{U}_p\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) \right| \boldsymbol{r}_\perp^{(\mathrm{i})} \right\rangle \\
&\quad \times \left\langle \boldsymbol{r}_\perp^{(\mathrm{i})} \left| \hat{U}^I_{(4)}\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) \right| \boldsymbol{r}_{\perp,\mathrm{o}} \right\rangle \psi(\boldsymbol{r}_{\perp,\mathrm{o}}, z_{\mathrm{o}})
\end{aligned}$$

$$
\begin{aligned}
&\sim \frac{1}{M}\exp\left(-\frac{\mathrm{i}\pi r_{\perp,\mathrm{i}}^2}{\lambda_0 f M}\right) \\
&\qquad \times \int d^2r_{\mathrm{o}} \int d^2r^{(\mathrm{i})}\delta^2\left(\boldsymbol{r}_{\perp}^{(\mathrm{i})} - \boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M\right) \\
&\qquad \times \left\langle \boldsymbol{r}_{\perp}^{(\mathrm{i})} \left| \hat{U}_{(4)}^{I}\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) \right| \boldsymbol{r}_{\perp,\mathrm{o}} \right\rangle \psi(\boldsymbol{r}_{\perp,\mathrm{o}}, z_{\mathrm{o}}) \\
&= \frac{1}{M}\exp\left(-\frac{\mathrm{i}\pi r_{\perp,\mathrm{i}}^2}{\lambda_0 f M}\right) \\
&\qquad \times \int d^2r_{\mathrm{o}} \left\langle \boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M \left| \hat{U}_{(4)}^{I}\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) \right| \boldsymbol{r}_{\perp,\mathrm{o}} \right\rangle \\
&\qquad \times \psi(\boldsymbol{r}_{\perp,\mathrm{o}}, z_{\mathrm{o}}) .
\end{aligned}
\tag{2.221}
$$

When there are no aberrations $\left\langle \boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M \left| \hat{U}_{(4)}^{I}\left(z_{\mathrm{i}}, z_{\mathrm{o}}\right) \right| \boldsymbol{r}_{\perp,\mathrm{o}} \right\rangle = \delta^2\left(\boldsymbol{r}_{\perp,\mathrm{o}} - \boldsymbol{r}_{\perp,\mathrm{i}}(\vartheta)/M\right)$ and hence one has the stigmatic imaging as seen earlier. It is clear from 2.221 that when aberrations are present the resultant intensity distribution in the image plane will represent only a blurred and distorted version of the image.

Usually, $\hat{U}_{(4)}^{I}$ is approximated by keeping only the most dominant aberration term, namely, the spherical aberration term which is independent of the position of the object-point. An important result to be recalled in this connection is the Scherzer's theorem [45, 46] which shows that the spherical aberration coefficient C_s is always positive and cannot be reduced below some minimum value governed by practical limitations.

By approximating $\hat{\mathcal{H}}_o$ by $\left(\hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,(4)}\right)$ we have only included the leading order aberrations called the third order aberrations. We are yet to take into account the contributions from $\hat{\mathcal{H}}_o^{(\lambda_0)}$ which are of order λ_0^2. We can believe that the effects of such terms are very small and their contributions can be likewise computed by using the interaction picture. The corrections to the paraxial map $\hat{U}_p\left(z, z_o\right)$ will come from the quadratic part of $\hat{\mathcal{H}}_o^{(\lambda_0)}$ which is of order λ_0^2. The term term $\hat{\mathbf{H}}_{o,(4)}^{(\lambda_0)}$ has to be added to $\hat{\mathbf{H}}_{o,(4)}$ to compute the corresponding $\hat{U}_{(4)}^{I}$ and this leads to the modification of the aberration coefficients. For example, the modified spherical aberration coefficient turns out be

$$\tilde{C}_s = \frac{1}{2}\int_{z_0}^{z_1} dz \left\{ \left(\alpha^4 - \alpha\alpha''\right) h^4 + \frac{\lambda_0^4}{8\pi^2}\left(\alpha\alpha''\right)' h^3 h' + 2\alpha^2 h^2 h'^2 + h'^4 \right\} . \qquad (2.222)$$

Since the nonclassical λ_0-dependent contribution to $\tilde{C}_s$ is very small compared to the dominant classical part Scherzer's theorem should not be affected though the above mentioned minimum value may change a little from the quantum contributions.

2.2.3 Magnetic Quadrupole Lens

In this case, we have

$$\begin{aligned} B &= \left(-Q_m y, -Q_m x, 0\right) , \\ Q_m &= \begin{cases} \text{constant in the lens region} & (z_l \leq z \leq z_r) \\ 0 \ \text{ outside the lens region} & (z < z_l \, , \, z > z_r) \end{cases} \end{aligned} \qquad (2.223)$$

corresponding to the vector potential

$$A = \left(0, 0, \frac{1}{2} Q_m \left(x^2 - y^2\right)\right) . \qquad (2.224)$$

Since there is no electric field in the lens region we can take $\phi(\boldsymbol{r}) = 0$. Then, from (2.62) the optical Hamiltonian $\hat{\mathcal{H}}_o$ is obtained as

$$\hat{\mathcal{H}}_o = -p_0 + \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,a} + \hat{\mathcal{H}}_o^{(\lambda_0)} \qquad (2.225)$$

$$\hat{\mathbf{H}}_{o,p} = \frac{1}{2p_0}\hat{p}_\perp^2 - \frac{1}{2} q Q_m \left(x^2 - y^2\right) \qquad (2.226)$$

$$\hat{\mathbf{H}}_{o,a} \approx \frac{1}{8p_0^3}\hat{p}_\perp^4 \qquad (2.227)$$

$$\hat{\mathcal{H}}_o^{(\lambda_0)} \approx \frac{\lambda_0^2 q Q_m}{16\pi^2 p_0^2}\left(\hat{p}_x^2 - \hat{p}_y^2\right) + \frac{\mathrm{i}\lambda_0 q Q_m}{4\pi p_0}\left(x\hat{p}_x - y\hat{p}_y\right) . \qquad (2.228)$$

Since $\hat{\mathbf{H}}_{o,p}$ is independent of z, the exact expression for the unitary paraxial transfer operator can be immediately written down: with $\Delta z = (z - z_0)$,

$$\hat{U}_p\left(z, z_0\right) = \mathrm{e}^{\frac{i}{\hbar} p_0 \Delta z} \exp\left\{ -\frac{\mathrm{i}}{\hbar}\left(\frac{\Delta z}{2p_0}\hat{p}_\perp^2 - \frac{1}{2} q Q_m \Delta z \left(x^2 - y^2\right)\right)\right\} . \qquad (2.229)$$

Taking $z_o < z_l$ and $z > z_r$ we have

$$
\begin{aligned}
\hat{U}_p(z, z_o) &= \hat{U}_{D,p}(z, z_r)\hat{U}_{L,p}(z_r, z_l)\hat{U}_{D,p}(z_l, z_o) \\
\hat{U}_{D,p}(z, z_r) &= e^{-\frac{i}{\hbar}(z-z_r)\left(-p_0+\frac{1}{2p_0}\hat{p}_\perp^2\right)} \\
\hat{U}_{L,p}(z_r, z_l) &= e^{-\frac{i}{\hbar}w\left(-p_0+\frac{1}{2p_0}\hat{p}_\perp^2-\frac{1}{2}qQ_m(x^2-y^2)\right)}, \quad w = (z_r - z_l) \\
\hat{U}_{D,p}(z_l, z_o) &= e^{-\frac{i}{\hbar}(z_l-z_o)\left(-p_0+\frac{1}{2p_0}\hat{p}_\perp^2\right)},
\end{aligned}
\tag{2.230}
$$

analogous to (2.165) in the case of the round lens. The corresponding paraxial transfer map for $(\boldsymbol{r}_\perp, \boldsymbol{p}_\perp)$ becomes

$$
\begin{aligned}
&\begin{pmatrix} \langle x\rangle(z) \\ \langle p_x\rangle(z)/p_0 \\ \langle y\rangle(z) \\ \langle p_y\rangle(z)/p_0 \end{pmatrix} = \begin{pmatrix} T_x & \mathrm{O} \\ \mathrm{O} & T_y \end{pmatrix} \begin{pmatrix} \langle x\rangle(z_o) \\ \langle p_x\rangle(z_o)/p_0 \\ \langle y\rangle(z_o) \\ \langle p_y\rangle(z_o)/p_0 \end{pmatrix} \\
&T_x = T_D(z - z_r)T_{xL}T_D(z_l - z_o) \\
&T_y = T_D(z - z_r)T_{yL}T_D(z_l - z_o) \\
&T_{xL} = \begin{pmatrix} \cosh\left(\sqrt{K}\,w\right) & \frac{1}{\sqrt{K}}\sinh\left(\sqrt{K}\,w\right) \\ \sqrt{K}\sinh\left(\sqrt{K}\,w\right) & \cosh\left(\sqrt{K}\,w\right) \end{pmatrix} \\
&T_{yL} = \begin{pmatrix} \cos\left(\sqrt{K}\,w\right) & \frac{1}{\sqrt{K}}\sin\left(\sqrt{K}\,w\right) \\ -\sqrt{K}\sin\left(\sqrt{K}\,w\right) & \cos\left(\sqrt{K}\,w\right) \end{pmatrix} \\
&\mathrm{O} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad T_D(d) = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, \quad K = \frac{qQ_m}{p_0}.
\end{aligned}
\tag{2.231}
$$

It is readily seen from this map that the lens is divergent (convergent) in the xz-plane and convergent (divergent) in yz-plane when $K > 0$ ($K < 0$). In other words, a line focus is produced by the quadrupole lens. In the weak field case, when $w^2 \ll 1/|K|$ (note that K has the dimension of (length)$^{-2}$) the lens can be considered as a thin lens with the focal lengths given by

$$
\frac{1}{f^{(x)}} = -\frac{1}{f^{(y)}} \approx -wK\,. \tag{2.232}
$$

Study of deviations from the ideal behaviour (2.231) due to $\mathbf{H}_{o,a}$ and $\hat{\mathcal{H}}_o^{(\lambda_0)}$ is straightforward using the same scheme employed above in the case of the magnetic round lens and we shall not consider it here.

In the field of electron optical technology, for particle energies in the range of tens or hundreds of kilovolts up to a few megavolts, quadrupole lens are used, if at all, as components in aberration-correcting units for round lenses and in devices required to produce a line focus. Quadrupole lenses are strong focusing : their fields exert a force directly on the electrons, towards or away from the axis, whereas in round magnetic lenses, the focusing force is more indirect, arising from the coupling between B_z and the azimuthal component of the electron velocity. So, it is mainly at higher energies, where round lenses are too weak, the strong focusing quadrupole lenses are exploited to provide the principal focusing field (see [1], for more details). Magnetic quadrupole lenses are the main components in beam transport systems in particle accelerators (for details see, *e.g.*, [47] and the recent text books, [3], [48], [49] and references therein).

2.2.4 Axially Symmetric Electrostatic Lens

An electrostatic round lens, with axis along the z-direction, comprises the electric field corresponding to the potential

$$\begin{aligned}\phi(\boldsymbol{r}_\perp, z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2 4^n} \phi^{(2n)}(z) \boldsymbol{r}_\perp^{2n} \\ &= \phi(z) - \frac{1}{4}\phi''(z) r_\perp^2 + \frac{1}{64}\phi''''(z) r_\perp^4 - \ldots, \end{aligned} \tag{2.233}$$

inside the lens region ($z_l \leq z \leq z_r$). Outside the lens $\phi = 0$. Using this value of $\phi(\boldsymbol{r})$ in (2.21) and (2.62), with $\boldsymbol{A} = (0,0,0)$, the beam-optical Hamiltonian of the lens takes the form,

$$\hat{\mathcal{H}}_o = -p_0 + \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,a} + \hat{\mathcal{H}}_o^{(\lambda_0)} \tag{2.234}$$

$$\begin{aligned}\hat{\mathbf{H}}_{o,p} &\approx 2p_0 (\eta - \mu(z)) \mu(z) \\ &+ \frac{1}{2p_0} \{1 + 2(\eta - \mu(z)) \mu(z)\} \hat{p}_\perp^2\end{aligned}$$

$$+\frac{1}{2}p_0\left\{1+2\left(\eta-\mu(z)\right)\mu(z)\right\}\left(2\mu(z)-\eta\right)\mu''(z)r_\perp^2$$

$$(2.235)$$

$$\begin{aligned}
\hat{\mathbf{H}}_{o,a} \approx\ & \frac{1}{8p_0^3}\hat{p}_\perp^4 \\
& +\frac{1}{32}p_0\left\{\left[4\left(\eta^2-1\right)-24\eta\mu(z)+24\mu(z)^2\right]\mu''(z)^2\right. \\
& \left.+\left[\eta+2\left(\eta^2-1\right)\mu(z)-6\eta\mu(z)^2+4\mu(z)^3\right]\mu^{(4)}\right\}r_\perp^4 \\
& +\frac{1}{8p_0}\left(2\mu(z)-\eta\right)\mu''(z)\left(r_\perp^2\hat{p}_\perp^2+\hat{p}_\perp^2r_\perp^2\right) \qquad (2.236)
\end{aligned}$$

$$\hat{\mathcal{H}}_o^{(\lambda_0)} \approx \frac{\lambda_0^2}{32\pi^2}(2\mu(z)-\eta)\mu''(z)\{\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp+\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp\} \qquad (2.237)$$

$$\eta=\frac{E+m_0c^2}{cp_0}, \qquad \mu(z)=\frac{q\phi(z)}{2cp_0}. \qquad (2.238)$$

The unitary paraxial transfer operator $\hat{U}_p(z,z_0)$ can be obtained as outlined in Appendix F, in terms of $\hat{\mathbf{H}}_{o,p}$ minus the first term $(-p_0)$ which contributes only a multiplicative phase factor to the wavefunction. In this case, unlike for the magnetic round lens, the coefficient of $\hat{p}_\perp^2$ is seen to depend on z. The calculation is straightforward and the paraxial transfer map reproduces the well-known classical results (see [1]). Here we have just demonstrated that $\hat{\mathcal{H}}_o$ can be brought to the general form for application of the general scheme of calculation of aberrations employed in the case of the magnetic round lens.

It may be noted that we have assumed the lens potential $\phi(\boldsymbol{r}_\perp,z)$ to vanish outside the lens region. In other words, we have considered the unipotential (einzel) lens having the same constant potential at both the object and the image side. There is no loss of generality in this assumption of our scheme since the so-called immersion lens, with two different constant potentials at the object and the image sides can also be treated using the same scheme simply by considering the right boundary (z_r) of the lens to be removed to infinity and including the constant value of the potential on the image side in the definition of $\phi(\boldsymbol{r}_\perp,z)$.

2.2.5 Electrostatic Quadrupole Lens

For the ideal electrostatic quadrupole lens with

$$\begin{aligned} \phi(\boldsymbol{r}) &= \frac{1}{2}Q_e\left(x^2 - y^2\right) , \\ Q_e &= \begin{cases} \text{constant in the lens region} & (z_l \leq z \leq z_r) \\ 0 \ \text{ outside the lens region} & (z < z_l\,,\ z > z_r)\,, \end{cases} \end{aligned} \tag{2.239}$$

and $\boldsymbol{A} = (0, 0, 0)$,

$$\hat{\mathcal{H}}_o = -p_0 + \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,a} \tag{2.240}$$

$$\hat{\mathbf{H}}_{o,p} = \frac{1}{2p_0}\hat{p}_\perp^2 + \frac{1}{2}p_0\eta\zeta\left(x^2 - y^2\right) \tag{2.241}$$

$$\begin{aligned} \hat{\mathbf{H}}_{o,a} \approx{}& \frac{1}{8p_0^3}\hat{p}_\perp^4 + \frac{1}{8}p_0\left(\eta^2 - 1\right)\zeta^2\left(x^2 - y^2\right)^2 \\ &+\frac{1}{8p_0}\eta\zeta\left\{\hat{p}_\perp^2\left(x^2 - y^2\right) + \left(x^2 - y^2\right)\hat{p}_\perp^2\right\} \end{aligned} \tag{2.242}$$

$$\eta = \frac{E + m_0c^2}{cp_0}\,, \qquad \zeta = \frac{qQ_e}{cp_0}\,, \tag{2.243}$$

and there are no λ_0-dependent terms up to this approximation. Simply by comparing $\hat{\mathbf{H}}_{o,p}$ in (2.241) with the $\hat{\mathbf{H}}_{o,p}$ of the magnetic quadrupole lens (2.226) it is immediately seen that a thin electrostatic quadrupole lens, of thickness $w = z_r - z_l$, has focal lengths given by

$$\frac{1}{f^{(x)}} = -\frac{1}{f^{(y)}} = \frac{wqQ_e\left(E + m_0c^2\right)}{c^2p_0^2} \tag{2.244}$$

Again, it is straightforward to study the deviations from the ideal behaviour using our general scheme.

Chapter 3

Spinor theory of charged-particle beam optics

3.1 Formalism

In this chapter we study the transport of the spin-$\frac{1}{2}$ particles through electromagnetic lenses, based on the Dirac equation, the basic equation for the spin-$\frac{1}{2}$ particles, taking fully into account the spinor character of the wavefunction. Such an approach based on the Dirac equation has been initiated by Jagannathan *et al.*, ([8]-[13]). The general formalism of the spinor theory will be illustrated through the examples of free propagatiion and round and quadrupole magnetic lenses.

Disregarding the anomalous magnetic moment (to be taken into account in Chapter IV), the quantum mechanics of a particle of mass m_0 and charge q moving in a static electromagnetic field with potentials $(\phi(\boldsymbol{r}), \boldsymbol{A}(\boldsymbol{r}))$ is governed by the time-dependent Dirac equation, written in the dimensionless form

$$\frac{i\hbar}{m_0c^2}\frac{\partial}{\partial t}\left(\begin{array}{c}\Psi_u(\boldsymbol{r},t)\\ \Psi_l(\boldsymbol{r},t)\end{array}\right) = \hat{\mathrm{H}}_D\left(\begin{array}{c}\Psi_u(\boldsymbol{r},t)\\ \Psi_l(\boldsymbol{r},t)\end{array}\right) \tag{3.1}$$

$$\Psi_u(\boldsymbol{r},t) = \left(\begin{array}{c}\Psi_1(\boldsymbol{r},t)\\ \Psi_2(\boldsymbol{r},t)\end{array}\right), \quad \Psi_l(\boldsymbol{r},t) = \left(\begin{array}{c}\Psi_3(\boldsymbol{r},t)\\ \Psi_4(\boldsymbol{r},t)\end{array}\right) \tag{3.2}$$

$$\hat{\mathrm{H}}_D = \beta + \hat{\mathcal{E}}_D + \hat{\mathcal{O}}_D\,, \quad \hat{\mathcal{E}}_D = \frac{q\phi(\boldsymbol{r})}{m_0c^2}\,, \quad \hat{\mathcal{O}}_D = \frac{\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}}}{m_oc} \tag{3.3}$$

$$\alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.4}$$

where Ψ_u an Ψ_l are the upper and lower components of the of the four-component Dirac spinor Ψ. The Dirac Hamiltonian in the above notation can be seen to be partitioned (apart from the leading term β) into an 'even part' $\hat{\mathcal{E}}_D$ and an 'odd part' $\hat{\mathcal{O}}_D$ with the algebraic property

$$\left[\beta, \hat{\mathcal{E}}_D\right] = 0, \qquad \left[\beta, \hat{\mathcal{O}}_D\right] = 2\beta\hat{\mathcal{O}}_D \tag{3.5}$$

As in the earlier chapters we shall follow the route of transforming the standard equations (Dirac equation this time) into a beam-optical form. As before, the study is confined to systems with straight optic axis chosen to be along the z-axis.

We are studying the action of the electromagnetic lens (situated between the planes $z = z_l$ and $z = z_r$) on an almost paraxial and quasimonoenergetic beam of Dirac particles being transported along the $+z$-direction. Under these conditions the spinor wavefunctions obeying the Dirac equation (3.1) take the form

$$\begin{aligned}
&\Psi(\boldsymbol{r}, t) \\
&\quad = \int_{p_0-\Delta p}^{p_0+\Delta p} dp\, \psi(\boldsymbol{r}; p) \exp\left(-\frac{i}{\hbar}\mathsf{E}(p)t\right) \\
&\quad p = |\boldsymbol{p}|, \quad \Delta p \ll p_0 \\
&\quad \mathsf{E}(p) = +\sqrt{m_0^2c^4 + c^2p^2} &(3.6) \\
&\psi(\boldsymbol{r}_\perp, z < z_l; p) \\
&\quad = \frac{1}{(2\pi\hbar)^{3/2}} \int\!\!\int dp_x\, dp_y\, \mathcal{U}(\boldsymbol{p}) \exp\left\{\frac{i}{\hbar}\left(\boldsymbol{p}_\perp \cdot \boldsymbol{r}_\perp + p_z z\right)\right\} \\
&\quad |\boldsymbol{p}_\perp| \ll p \\
&\quad \boldsymbol{p} = \left(\boldsymbol{p}_\perp, p_z = +\sqrt{p^2 - p_\perp^2}\right) &(3.7) \\
&\mathcal{U}(\boldsymbol{p}) = a_+(\boldsymbol{p})u_+(\boldsymbol{p}) + a_-(\boldsymbol{p})u_-(\boldsymbol{p})
\end{aligned}$$

$$
= \sqrt{\frac{\mathsf{E}(p) + m_0c^2}{2\mathsf{E}(p)}} \begin{pmatrix} a_+(\boldsymbol{p}) \\ a_-(\boldsymbol{p}) \\ \frac{c[a_+(\boldsymbol{p})p_z + a_-(\boldsymbol{p})(p_x - \mathrm{i}p_y)]}{\mathsf{E}(p)+m_0c^2} \\ \frac{c[a_+(\boldsymbol{p})(p_x + \mathrm{i}p_y) - a_-(\boldsymbol{p})p_z]}{\mathsf{E}(p)+m_0c^2} \end{pmatrix}
$$

$$
\int d^3p \left(|a_+(\boldsymbol{p})|^2 + |a_-(\boldsymbol{p})|^2 \right) = 1 \,, \tag{3.8}
$$

where $\left\{u_\pm(\boldsymbol{p}) \exp\left[\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r} - \mathsf{E}(p)t)\right]\right\}$ are the standard positive-energy plane-wave solutions of the free-particle Dirac equation. We are interested in obtaining a relation for the wavefunctions at different planes along the z-axis. So, we assume a relationship of the type

$$
\begin{aligned}
&\psi(\boldsymbol{r}_\perp, z^{(2)}; p) \\
&= \sum_k \int d^2r^{(1)} \left\langle \boldsymbol{r}_\perp^{(2)} \left| \hat{\mathcal{T}}_{jk}\left(z^{(2)}, z^{(1)}; p\right) \right| \boldsymbol{r}_\perp^{(1)} \right\rangle \psi_k(\boldsymbol{r}_\perp^{(1)}, z^{(1)}; p) \\
&\qquad j, k = 1, 2, 3, 4.,
\end{aligned} \tag{3.9}
$$

for $\psi(\boldsymbol{r}_\perp, z; p)$. Then we have

$$
\begin{aligned}
&\left|\Psi\left(z^{(2)}, t\right)\right\rangle \\
&= \int_{p_0-\Delta p}^{p_0+\Delta p} dp \, \exp\left\{-\frac{\mathrm{i}}{\hbar}\mathsf{E}(p)t\right\} \hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}; p\right) \left|\psi\left(z^{(1)}; p\right)\right\rangle \\
&\approx \hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}; p_0\right) \left|\Psi\left(z^{(1)}; t\right)\right\rangle \,, \\
&\qquad \text{in the practically monoenergetic case } (\Delta p \approx 0) \,.
\end{aligned} \tag{3.10}
$$

Since we are taking the beam to be practically monoenergetic and assigning a mean momentum p_0 to the incident beam particle, the wavefunction takes the form

$$
\Psi(\boldsymbol{r}, t) = \mathrm{e}^{-\mathrm{i}\mathsf{E}t/\hbar} \psi\left(\boldsymbol{r}_\perp, z; p_0\right) \,. \tag{3.11}
$$

This leads to the time-independent equation for $\psi\left(\boldsymbol{r}_\perp, z; p_0\right)$ given by

$$
\begin{aligned}
&\left\{ \frac{\mathsf{E}\left(p_0\right)}{m_0c^2} - \beta - \frac{q\phi}{m_0c^2} - \frac{\boldsymbol{\alpha}_\perp \cdot \hat{\boldsymbol{\pi}}_\perp}{m_0c} \right. \\
&\qquad \left. + \frac{1}{m_0c}\alpha_z \left(\mathrm{i}\hbar\frac{\partial}{\partial z} + qA_z \right) \right\} \psi\left(\boldsymbol{r}_\perp, z; p_0\right) = 0 \,.
\end{aligned} \tag{3.12}
$$

Equation (3.12) is already linear in $\frac{\partial}{\partial z}$ as desired. The required z-propagator $\hat{\mathcal{T}}\left(z^{(2)}, z^{(1)}; p_0\right)$ is to be obtained by integrating (3.12).

To proceed further, the time-independent Dirac equation (3.12) has to be cast into a suitable optical form. To this end, we start by multiplying (3.12) by $\frac{m_0 c \alpha_z}{p_0}$ throughout from the left and rearranging the terms. We get

$$\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial}{\partial z}\psi\left(\boldsymbol{r}_\perp, z; p_0\right) = \hat{H}\psi\left(\boldsymbol{r}_\perp, z; p_0\right) \tag{3.13}$$

$$\hat{H} = -\beta\chi\alpha_z + \frac{q\phi}{cp_0}\alpha_z + \frac{1}{p_0}\alpha_z\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp - \frac{q}{p_0}A_z \tag{3.14}$$

$$\chi = \begin{pmatrix} \xi\mathbb{1} & \mathbf{0} \\ \mathbf{0} & -\xi^{-1}\mathbb{1} \end{pmatrix} = \frac{1}{cp_0}\begin{pmatrix} \left(E + 2m_0c^2\right)\mathbb{1} & \mathbf{0} \\ \mathbf{0} & -E\mathbb{1} \end{pmatrix}, \tag{3.15}$$

where E is the kinetic energy of the incident beam particle. The next and the crucial step lies in making use of the transformation [8]

$$\psi \longrightarrow \psi' = M\psi\,, \quad M = \frac{1}{\sqrt{2}}\left(I + \chi\alpha_z\right), \quad M^{-1} = \frac{1}{\sqrt{2}}\left(I - \chi\alpha_z\right). \tag{3.16}$$

where I is the 4×4 identity matrix. Then, ψ' satisfies

$$\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi'}{\partial z} = \left(M\hat{H}M^{-1}\right)\psi' = \hat{H}'\psi' \tag{3.17}$$

$$\hat{H}' = -\beta + \hat{\mathcal{E}} + \hat{\mathcal{O}} \tag{3.18}$$

$$\hat{\mathcal{E}} = \frac{\eta q\phi}{cp_0}\beta - \frac{1}{p_0}qA_z \tag{3.19}$$

$$\hat{\mathcal{O}} = \frac{1}{p_0}\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp - \frac{1}{p_0^2}m_0q\phi\chi\alpha_z \tag{3.20}$$

$$\eta = \frac{E + m_0c^2}{cp_0}. \tag{3.21}$$

Equation (3.17) is the desired beam-optical representation of the time-independent Dirac equation, with a striking similarity to the standard time-dependent Dirac equation: $\hat{H}' = -\beta +$ an odd part $+$ an even part . However it is to be pointed out that there is a difference that $\hat{H}'$ is not hermitian, unlike the standard Dirac Hamiltonian (3.3). Physically this means that $\sum_{j=1}^{4}\int d^2r\,|\psi_j(\boldsymbol{r}_\perp, z)|^2$, the probability of finding the beam particle in the xy-plane need not be a constant, in general, along the z-axis.. Only the total probability for the existence of the particle in the entire space is conserved in the absence of any mechanism for particle creation and annihilation.

The next step lies in exploiting the above mentioned similarity. This is done by employing a Foldy-Wouthuysen-like transformation technique to expand the beam-optical Hamiltonian $\hat{H}'$ into the paraxial and aberrating terms, analogous to the way the usual Foldy-Wouthuysen transformation expands the standard Dirac Hamiltonian into the nonrelativistic part and relativistic corrections part, and filtering out the part of the z-evolution equation relevant for the beam propagating in the $+z$-direction. Then, using the standard techniques of quantum theory for studying the time-evolution, the z-evolution of the spinor wavefunction of the forward propagating beam is studied up to any desired level of accuracy.

Note that we are dealing with an almost monoenergetic quasiparaxial incident beam with design momentum $p_0 = |\boldsymbol{p}|$, $p_z > 0$, $|\boldsymbol{p}_\perp| \ll p_z$ and $p_z = \sqrt{p_0^2 - p_\perp^2} \approx p_0$. We also make note that the strength of the odd operator $\hat{\mathcal{O}}$ in $\hat{H}'$ is of the order of $\frac{1}{p_0}$. Using $\frac{1}{p_0}$ as the expansion parameter the first transformation

$$\psi' \longrightarrow \psi^{(1)} = \mathrm{e}^{\mathrm{i}\hat{S}_1}\psi', \quad \hat{S}_1 = \frac{\mathrm{i}}{2}\beta\hat{\mathcal{O}}\,, \tag{3.22}$$

leads to the result

$$\begin{aligned}
\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi'}{\partial z} &= \hat{H}^{(1)}\psi^{(1)} && (3.23)\\
\hat{H}^{(1)} &= -\beta + \hat{\mathcal{E}}_1 + \hat{\mathcal{O}}_1 && (3.24)\\
\hat{\mathcal{E}}_1 &\approx \hat{\mathcal{E}} - \frac{1}{2}\beta\hat{\mathcal{O}}^2 - \frac{1}{8}\left[\hat{\mathcal{O}}, \left(\left[\hat{\mathcal{O}}, \hat{\mathcal{E}}\right] + \frac{\mathrm{i}\lambda_0}{2\pi}\left(\frac{\partial\hat{\mathcal{O}}}{\partial z}\right)\right)\right] \\
&\quad + \frac{1}{8}\hat{\mathcal{O}}^4 && (3.25)\\
\hat{\mathcal{O}}_1 &\approx -\frac{1}{2}\beta\left(\left[\hat{\mathcal{O}}, \hat{\mathcal{E}}\right] + \frac{\mathrm{i}\lambda_0}{2\pi}\left(\frac{\partial\hat{\mathcal{O}}}{\partial z}\right)\right) - \frac{1}{3}\hat{\mathcal{O}}^3\,. && (3.26)
\end{aligned}$$

It is seen that the first transformation (3.22) has reduced the strength of the odd operator $\hat{\mathcal{O}}_1$ to the order of $\frac{1}{p_0^2}$. It is also to be noted that the transformation in (3.22) is not unitary.

To reduce the strength of the odd operator further we make a second transformation

$$\psi^{(1)} \longrightarrow \psi^{(2)} = \mathrm{e}^{\mathrm{i}\hat{S}_2}\psi^{(1)}\,, \quad \hat{S}_2 = \frac{\mathrm{i}}{2}\beta\hat{\mathcal{O}}_1\,, \tag{3.27}$$

which gives

$$\begin{aligned}
\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi^{(2)}}{\partial z} &= \hat{H}^{(2)}\psi^{(2)} && (3.28)\\
\hat{H}^{(2)} &= -\beta + \hat{\mathcal{E}}_2 + \hat{\mathcal{O}}_2 && (3.29)\\
\hat{\mathcal{E}}_2 &= \hat{\mathcal{E}}_1\left(\hat{\mathcal{E}} \to \hat{\mathcal{E}}_1\,, \hat{\mathcal{O}} \to \hat{\mathcal{O}}_1\right) && (3.30)\\
\hat{\mathcal{O}}_2 &= \hat{\mathcal{O}}_1\left(\hat{\mathcal{E}} \to \hat{\mathcal{E}}_1\,, \hat{\mathcal{O}} \to \hat{\mathcal{O}}_1\right)\,, && (3.31)
\end{aligned}$$

Now, the strength of the odd operator $\hat{\mathcal{O}}_2$ is of order $\frac{1}{p_0^3}$. After another such transformation

$$\psi^{(2)} \longrightarrow \psi^{(3)} = \mathrm{e}^{\mathrm{i}\hat{S}_3}\psi^{(2)}\,, \quad \hat{S}_3 = \frac{\mathrm{i}}{2}\beta\hat{\mathcal{O}}_2\,, \tag{3.32}$$

we have

$$\begin{aligned}
\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi^{(3)}}{\partial z} &= \hat{H}^{(3)}\psi^{(3)} && (3.33)\\
\hat{H}^{(3)} &= -\beta + \hat{\mathcal{E}}_3 + \hat{\mathcal{O}}_3 && (3.34)\\
\hat{\mathcal{E}}_3 &= \hat{\mathcal{E}}_1\left(\hat{\mathcal{E}} \to \hat{\mathcal{E}}_2\,, \hat{\mathcal{O}} \to \hat{\mathcal{O}}_2\right) && (3.35)\\
\hat{\mathcal{O}}_3 &= \hat{\mathcal{O}}_1\left(\hat{\mathcal{E}} \to \hat{\mathcal{E}}_2\,, \hat{\mathcal{O}} \to \hat{\mathcal{O}}_2\right)\,, && (3.36)
\end{aligned}$$

with the odd operator $\hat{\mathcal{O}}_3$ of the order of $\frac{1}{p_0^4}$. Thus, having gone up to the required order of accuracy, which is sufficient for working out the third order aberrations, we drop from $\hat{H}^{(3)}$ the odd operator $\hat{\mathcal{O}}_3$ and write

$$\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi^{(3)}}{\partial z} \approx \hat{H}^{(3)}\psi^{(3)} \tag{3.37}$$

$$\begin{aligned} \hat{H}^{(3)} &= -\beta + \hat{\mathcal{E}} - \frac{1}{2}\beta\hat{\mathcal{O}}^2 - \frac{1}{8}\left[\hat{\mathcal{O}}, \left(\left[\hat{\mathcal{O}}, \hat{\mathcal{E}}\right] + \frac{\mathrm{i}\lambda_0}{2\pi}\left(\frac{\partial\hat{\mathcal{O}}}{\partial z}\right)\right)\right] \\ &\quad + \frac{1}{8}\beta\left\{\hat{\mathcal{O}}^4 + \left(\left[\hat{\mathcal{O}}, \hat{\mathcal{E}}\right] + \frac{\mathrm{i}\lambda_0}{2\pi}\left(\frac{\partial\hat{\mathcal{O}}}{\partial z}\right)\right)^2\right\} . \end{aligned} \tag{3.38}$$

Let us examine the the expression for $\hat{H}^{(3)}$ in (3.38). In general, it can always be written in the form,

$$\hat{H}^{(3)} = \begin{pmatrix} \hat{h}_1 & \mathbf{O} \\ \mathbf{O} & \hat{h}_1 \end{pmatrix} + \beta \begin{pmatrix} \hat{h}_2 & \mathbf{O} \\ \mathbf{O} & \hat{h}_2 \end{pmatrix} = \begin{pmatrix} \hat{h}_1 + \hat{h}_2 & \mathbf{O} \\ \mathbf{O} & \hat{h}_1 - \hat{h}_2 \end{pmatrix} . \tag{3.39}$$

We also know that the lower components of $\psi^{(3)}$ in (3.37) are small compared to the upper components. So we can write

$$\beta\psi^{(3)} \approx \psi^{(3)} . \tag{3.40}$$

Naively speaking we can say that we can drop β from $\hat{H}^{(3)}$ written as in (3.39). With this approximation, equation (3.37) takes the form

$$\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi^{(3)}}{\partial z} \approx \hat{\mathsf{H}}\psi^{(3)} , \qquad \hat{\mathsf{H}} = \begin{pmatrix} \hat{h}_1 + \hat{h}_2 & \mathbf{O} \\ \mathbf{O} & \hat{h}_1 + \hat{h}_2 \end{pmatrix} . \tag{3.41}$$

It is to be noted that the expression obtained for $\hat{\mathsf{H}}$ is in a Foldy-Wouthuysen representation. In this chapter we are primarily interested in the problem of imaging, for which we need to know the wavefunction at the image plane correspond to the wavefunction at the object plane. In order to enable a direct interpretation in terms of the more familiar Dirac representation, we return back to the original Dirac representation. To get back to the original Dirac representation we have to retrace through the following transformations

$$\psi^{(3)} \longrightarrow \psi = M^{-1}\mathrm{e}^{-\mathrm{i}\hat{S}_1}\mathrm{e}^{-\mathrm{i}\hat{S}_2}\mathrm{e}^{-\mathrm{i}\hat{S}_3}\psi^{(3)} \approx M^{-1}\mathrm{e}^{-\mathrm{i}\hat{\mathcal{S}}}\psi^{(3)} \tag{3.42}$$

$$\begin{aligned} \mathcal{S} &\approx \hat{S}_1 + \hat{S}_2 + \hat{S}_3 - \frac{\mathrm{i}}{2}\left(\left[\hat{S}_1 , \hat{S}_2\right] + \left[\hat{S}_1 , \hat{S}_3\right]\right. \\ &\quad \left. + \left[\hat{S}_2 , \hat{S}_3\right]\right) - \frac{1}{4}\left[\left[\hat{S}_1 , \hat{S}_2\right] , \hat{S}_3\right] \ldots . \end{aligned} \tag{3.43}$$

Retaining the terms consistent to the same order of accuracy with which we are working we finally get

$$
\begin{aligned}
\mathrm{i}\hbar\frac{\partial}{\partial z}|\psi(z)\rangle &= \hat{\mathcal{H}}_o|\psi(z)\rangle && (3.44)\\
\hat{\mathcal{H}}_o &= M^{-1}\left\{\mathrm{e}^{-\mathrm{i}\hat{\mathcal{S}}}p_0\hat{\mathsf{H}}\mathrm{e}^{\mathrm{i}\hat{\mathcal{S}}} - \mathrm{i}\hbar\mathrm{e}^{-\mathrm{i}\hat{\mathcal{S}}}\frac{\partial}{\partial z}\left(\mathrm{e}^{\mathrm{i}\hat{\mathcal{S}}}\right)\right\}M \\
&= M^{-1}\left\{p_0\hat{\mathsf{H}} + \hbar\frac{\partial\hat{\mathcal{S}}}{\partial z} - \mathrm{i}\left[\hat{\mathcal{S}}\,,\,p_0\hat{\mathsf{H}} + \frac{1}{2}\hbar\frac{\partial\hat{\mathcal{S}}}{\partial z}\right]\right. \\
&\qquad -\frac{1}{2!}\left[\hat{\mathcal{S}}\,,\,\left[\hat{\mathcal{S}}\,,\,p_0\hat{\mathsf{H}} + \frac{1}{3}\hbar\frac{\partial\hat{\mathcal{S}}}{\partial z}\right]\right] \\
&\qquad \left. +\frac{\mathrm{i}}{3!}\left[\hat{\mathcal{S}}\,,\,\left[\hat{\mathcal{S}}\,,\,\left[\hat{\mathcal{S}}\,,\,p_0\hat{\mathsf{H}} + \frac{1}{4}\hbar\frac{\partial\hat{\mathcal{S}}}{\partial z}\right]\right]\right]\cdots\right\}M\,. && (3.45)
\end{aligned}
$$

The resulting beam-optical Hamiltonian of the Dirac particle can be written, in general, in the following form:

$$
\begin{aligned}
\hat{\mathcal{H}}_o &= -p_0 + \hat{\mathbf{H}}_{o,p} + \hat{\mathcal{H}}_{o,a} && (3.46)\\
\hat{\mathcal{H}}_{o,a} &= \hat{\mathbf{H}}_{o,a} + \hat{\mathcal{H}}_o^{(\lambda_0)} + \hat{\mathcal{H}}_o^{(\lambda_0,\sigma)}\,, && (3.47)
\end{aligned}
$$

where $\hat{\mathbf{H}}_{o,p}$, $\hat{\mathbf{H}}_{o,a}$ and $\hat{\mathcal{H}}_o^{(\lambda_0)}$ are scalar terms ($\sim I$) and $\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)}$ is a 4×4 matrix term which also vanishes in the limit $\lambda_0 \longrightarrow 0$, like $\hat{\mathcal{H}}_o^{(\lambda_0)}$. Now, the performance of the optical system under study, corresponding to the assumed values of the potentials $\phi(\boldsymbol{r})$ and $\boldsymbol{A}(\boldsymbol{r})$, can be calculated using the same scheme described in detail in Chapter II; the only difference is that in the Dirac case the Hamiltonian $\hat{\mathcal{H}}_o$ is a 4×4 matrix with operator entries and the wavefunction is a 4-component column vector. The matrix term $\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)}$ can be clubbed with $\hat{\mathbf{H}}_{o,a}$ and $\hat{\mathcal{H}}_o^{(\lambda_0)}$, as indicated above, and treated using the interaction picture. It is found that the beam-optical Hamiltonians $\left(\hat{\mathcal{H}}_o\right)$ in the Klein-Gordon theory and the Dirac theory do not differ in their classical parts $\left(\hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,a}\right)$. Thus the Klein-Gordon theory without the term $\hat{\mathcal{H}}_o^{(\lambda_0)}$ and the Dirac theory without the terms $\hat{\mathcal{H}}_o^{(\lambda_0)}$ and $\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)}$ are identical, effectively, as seen below.

Note that for an observable O of the Dirac particle, with the corresponding hermitian operator $\hat{O}$ given in a 4×4 matrix form, the expectation value is defined by

$$\begin{aligned} \langle O \rangle (z) &= \frac{\langle \psi(z) | \hat{O} | \psi(z) \rangle}{\langle \psi(z) | \psi(z) \rangle} \\ &= \frac{\sum_{j,k=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z) \hat{O}_{jk} \psi_k(\boldsymbol{r}_\perp, z)}{\sum_{j=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z) \psi_j(\boldsymbol{r}_\perp, z)} . \end{aligned} \quad (3.48)$$

Hence, the map $\langle O \rangle (z_o) \longrightarrow \langle O \rangle (z)$ becomes

$$\begin{aligned} \langle O \rangle (z) &= \frac{\langle \psi(z_o) | \hat{\mathcal{T}}^\dagger (z, z_o) \hat{O} \hat{\mathcal{T}} (z, z_o) | \psi(z_o) \rangle}{\langle \psi(z_o) | \hat{\mathcal{T}}^\dagger (z, z_o) \hat{\mathcal{T}} (z, z_o) \psi(z_o) \rangle} \\ &= \frac{\sum_{j,l,m,k=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z_o) \hat{\mathcal{T}}_{jl}^\dagger (z, z_o) \hat{O}_{lm} \hat{\mathcal{T}}_{mk} (z, z_o) \psi_k(\boldsymbol{r}_\perp, z)}{\sum_{j,m,k=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z_o) \hat{\mathcal{T}}_{jm}^\dagger (z, z_o) \hat{\mathcal{T}}_{mk} (z, z_o) \psi_k(\boldsymbol{r}_\perp, z_o)} . \end{aligned} \quad (3.49)$$

For position and momentum operators $(\sim I)$,

$$\begin{aligned} \langle \boldsymbol{r}_\perp \rangle (z) &= \\ &= \frac{\sum_{j,m,k=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z_o) \hat{\mathcal{T}}_{jm}^\dagger (z, z_o) \boldsymbol{r}_\perp \hat{\mathcal{T}}_{mk} (z, z_o) \psi_k(\boldsymbol{r}_\perp, z)}{\sum_{j,m,k=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z_o) \hat{\mathcal{T}}_{jm}^\dagger (z, z_o) \hat{\mathcal{T}}_{mk} (z, z_o) \psi_k(\boldsymbol{r}_\perp, z_o)} \end{aligned} \quad (3.50)$$

$$\begin{aligned} \langle \boldsymbol{p}_\perp \rangle (z) &= \\ &= \frac{\sum_{j,m,k=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z_o) \hat{\mathcal{T}}_{jm}^\dagger (z, z_o) \hat{\boldsymbol{p}}_\perp \hat{\mathcal{T}}_{mk} (z, z_o) \psi_k(\boldsymbol{r}_\perp, z)}{\sum_{j,m,k=1}^{4} \int d^2 r \psi_j^*(\boldsymbol{r}_\perp, z_o) \hat{\mathcal{T}}_{jm}^\dagger (z, z_o) \hat{\mathcal{T}}_{mk} (z, z_o) \psi_k(\boldsymbol{r}_\perp, z_o)} . \end{aligned} \quad (3.51)$$

When the terms $\hat{\mathcal{H}}_0^{(\lambda_0)}$ and $\hat{\mathcal{H}}_0^{(\lambda_0, \sigma)}$ are dropped from the Dirac optical Hamiltonian it becomes $\sim I$ and the corresponding transfer operator also becomes $\sim I$ with respect to the spinor index: i.e., $\hat{\mathcal{T}}_{jm}(z, z_o) = \hat{\mathcal{T}}(z, z_o) \delta_{jm}$. Then, though all the four components of ψ, $(\psi_1, \psi_2, \psi_3, \psi_4)$, contribute to the averages of $\boldsymbol{r}_\perp, \boldsymbol{p}_\perp$, etc., as seen from the above definitions, one can think of them as due to a single component ψ, effectively, since the contributions from the four components cannot be identified individually in the final results. Thus, in this case, there would be no difference between the 'classical' transfer map for $(\langle \boldsymbol{r}_\perp\rangle(z), \langle \boldsymbol{p}_\perp\rangle(z))$ ((2.209)–(2.212)) and the corresponding transfer map in the Dirac theory. In this sense, the Dirac theory and the Klein-Gordon theory are identical scalar theories when λ_0-dependent terms are ignored in the Klein-Gordon theory and λ_0-dependent scalar and matrix terms are ignored in the Dirac theory.

We shall consider below, very briefly, a few specific examples of the above formalism of the Dirac theory of charged-particle wave optics.

3.2 Applications

3.2.1 Free Propagation : Diffraction

For a monoenergetic quasiparaxial Dirac beam propagating in free space along the $+z$-direction equation (3.17) reads

$$\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial \psi'}{\partial z} = \hat{H}'\psi' , \qquad \hat{H}' = -\left(\beta - \frac{1}{p_0}\chi \boldsymbol{\alpha}_\perp \cdot \hat{\boldsymbol{p}}_\perp\right) , \tag{3.52}$$

with

$$\left(p_0 \hat{H}'\right)^2 = \left(p_0^2 - \hat{p}_\perp^2\right) I . \tag{3.53}$$

Thus, $p_0\hat{H}' = -p_0\beta + \chi\boldsymbol{\alpha}_\perp \cdot \hat{\boldsymbol{p}}_\perp$ can be identified with the classical optical Hamiltonian $-\sqrt{p_0^2 - \hat{p}_\perp^2}$, for free propagation of a monoenergetic quasiparaxial beam, with the square root taken in the Dirac way. Though in the present case it may look as if one can take such a square root using only the three 2×2 Pauli σ-matrices, it is necessary to use the 4×4 Dirac matrices in order to take into account the two-component spin and the propagations in the forward and backward directions along the z-axis considered separate. It can be verified that for the paraxial plane-wave solutions of (3.52) corresponding to forward propagation in the $+z$ direction, with $p_z > 0$ and $|\boldsymbol{p}_\perp| \ll p_z \approx p_0$, the upper pair of components are large compared to the lower pair of components, analogous to the nonrelativistic positive-energy solutions of the free-particle Dirac equation.

In the same way as the free-particle Dirac Hamiltonian can be diagonalized by a Foldy-Wouthuysen transformation (see Appendix C) the odd part in $\hat{H}'$ can be completely removed by a transformation: with

$$\psi'' = \mathrm{e}^{-\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{p}}_\perp\theta}\psi' , \qquad \tanh 2|\hat{\boldsymbol{p}}_\perp|\theta = \frac{|\hat{\boldsymbol{p}}_\perp|}{p_0} , \tag{3.54}$$

we have

$$\begin{aligned}
\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi''}{\partial z} &= \hat{H}''\psi'' && (3.55)\\
\hat{H}'' &= \mathrm{e}^{-\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{p}}_\perp\theta}\hat{H}'\mathrm{e}^{\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{p}}_\perp\theta} \\
&= \left(\cosh|\hat{\boldsymbol{p}}_\perp|\theta - \frac{\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{p}}_\perp}{|\hat{\boldsymbol{p}}_\perp|}\sinh|\hat{\boldsymbol{p}}_\perp|\theta\right)\hat{H}' \\
&\quad\times\left(\cosh|\hat{\boldsymbol{p}}_\perp|\theta - \frac{\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{p}}_\perp}{|\hat{\boldsymbol{p}}_\perp|}\sinh|\hat{\boldsymbol{p}}_\perp|\theta\right) \\
&= -\frac{1}{p_0}\left(\sqrt{p_0^2-\hat{p}_\perp^2}\right)\beta . && (3.56)
\end{aligned}$$

Now, invoking the fact that ψ'' will have lower components very small compared to the upper components in the quasiparaxial situation, we can write

$$\frac{\mathrm{i}\lambda_0}{2\pi}\frac{\partial\psi''}{\partial z} \approx -\frac{1}{p_0}\left(\sqrt{p_0^2-\hat{p}_\perp^2}\right)\psi'' . \tag{3.57}$$

Then, making the inverse transformation

$$\psi = M^{-1}\mathrm{e}^{\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{p}}_\perp\theta}\psi'' , \tag{3.58}$$

equation (3.57) becomes

$$\begin{aligned}
\mathrm{i}\hbar\frac{\partial\psi}{\partial z} &= \hat{\mathcal{H}}_o\psi && (3.59)\\
\hat{\mathcal{H}}_o &= -\left(\sqrt{p_0^2-\hat{p}_\perp^2}\right) \approx -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2 + \frac{1}{8p_0^3}\hat{p}_\perp^4 + \ldots , && (3.60)
\end{aligned}$$

exactly as in the scalar case (see (2.79)) except for the fact that now ψ has four components. Then, it is obvious that the diffraction pattern due to a quasiparaxial Dirac-particle beam will be the superposition of the patterns due to the four individual components $(\psi_1, \psi_2, \psi_3, \psi_4)$ of the spinor ψ representing the beam : for a highly paraxial beam the intensity distribution of the diffraction pattern at the xy-plane at z will be given by (see (2.86))

$$I(x,y,z) \sim \sum_{j=1}^{4}\left|\int\int dx_0\, dy_0 \exp\left\{\frac{\mathrm{i}p_0}{2\hbar(z-z_o)}\left[(x-x_0)^2 + (y-y_0)^2\right]\right\}\psi_j(x_0,y_0,z_o)\right|^2, \quad (3.61)$$

where the plane of the diffracting object is at z_o. It is clear that when the presence of a field makes $\hat{\mathcal{H}}_o$ acquire a matrix component ($\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)}$) the transfer operator $\hat{\mathcal{T}}(z,z_o)$ would have a nontrivial matrix structure leading to an interference between the diffracted amplitudes $(\psi_1,\psi_2,\psi_3,\psi_4)$.

When the monoenergetic beam is not sufficiently paraxial to allow the approximations made above one can directly use the free z-evolution equation

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial z} = -\left\{p_0\beta\chi\alpha_z + \mathrm{i}\left(\Sigma_x\hat{p}_y - \Sigma_y\hat{p}_x\right)\right\}\psi \quad (3.62)$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}, \quad (3.63)$$

obtained by setting $\phi = 0$ and $\boldsymbol{A} = (0,0,0)$ in equations (3.13)-(3.15). Integrating (3.62), we have

$$|\psi(z)\rangle = \exp\left\{\frac{\mathrm{i}}{\hbar}\Delta z\left(p_0\beta\chi\alpha_z + \mathrm{i}\left(\Sigma_x\hat{p}_y - \Sigma_y\hat{p}_x\right)\right)\right\}|\psi(z_o)\rangle$$

$$\Delta z = (z - z_o), \quad (3.64)$$

the general law of propagation of the free Dirac wavefunction in the $+z$-direction, showing the subtle way in which the Dirac equation mixes up the spinor components (for some detailed studies on the optics of general free Dirac waves, in particular, diffraction, see [50]-[60]).

3.2.2 Axially Symmetric Magnetic Lens

In this case, following the procedure of obtaining $\hat{\mathcal{H}}_o$ as outlined above, we get

$$\hat{\mathcal{H}}_o = -p_0 + \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,a} + \hat{\mathcal{H}}_o^{(\lambda_0)} + \hat{\mathcal{H}}_o^{(\lambda_0,\sigma)} \quad (3.65)$$

$$
\begin{aligned}
\hat{\mathbf{H}}_{o,p} &\approx \frac{1}{2p_0}\hat{p}_\perp^2 + \frac{1}{2}p_0\alpha^2(z)r_\perp^2 - \alpha(z)\hat{L}_z\,, & (3.66)\\
\hat{\mathbf{H}}_{o,a} &\approx \frac{1}{8p_0^3}\hat{p}_\perp^4 - \frac{1}{2p_0^2}\alpha(z)\hat{p}_\perp^2\hat{L}_z \\
&\quad -\frac{1}{8p_0}\alpha(z)^2\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp + \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)^2 \\
&\quad +\frac{3}{8p_0}\alpha(z)^2\left(\hat{p}_\perp^2 r_\perp^2 + r_\perp^2\hat{p}_\perp^2\right) \\
&\quad +\frac{1}{8}\left(\alpha''(z) - 4\alpha(z)^3\right)\hat{L}_z r_\perp^2 \\
&\quad +\frac{p_0}{8}\left(\alpha(z)^4 - \alpha(z)\alpha''(z)\right)r_\perp^4\,, & (3.67)\\
\hat{\mathcal{H}}_o^{(\lambda_0)} &\approx \frac{p_0\lambda_0^2}{32\pi^2}\left(\alpha'(z)^2 - 2\alpha(z)\alpha''(z)\right)r_\perp^2 \\
&\quad +\frac{p_0\lambda_0^2}{128\pi^2}\left(\alpha''(z)^2 - \alpha'(z)\alpha'''(z)\right)r_\perp^4 & (3.68)\\
\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)} &\approx \left\{\frac{\mathrm{i}\lambda_0^2}{32\pi^2}\alpha'(z)\hat{p}_x - \frac{\mathrm{i}p_0\lambda_0^2}{32\pi^2}\alpha(z)\alpha'(z)y + \frac{\mathrm{i}p_0\lambda_0^2}{16\pi^2}\alpha''(z)x\right. \\
&\quad -\frac{\mathrm{i}\lambda_0^2}{128\pi^2}\alpha'''(z)\hat{p}_x r_\perp^2 - \frac{\mathrm{i}\lambda_0^2}{128\pi^2}\alpha'''(z)x\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp\right. \\
&\quad \left.+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right) - \frac{\mathrm{i}p_0\lambda_0^2}{128\pi^2}\alpha^{(4)}(z)xr_\perp^2 \\
&\quad \left.+\frac{p_0\lambda_0^2}{256\pi^2}\left(3\alpha'(z)\alpha''(z) - \mathrm{i}\alpha(z)\alpha'''(z)\right)yr_\perp^2\right\}\Sigma_x \\
&\quad +\left\{\frac{\mathrm{i}\lambda_0^2}{32\pi^2}\alpha'(z)\hat{p}_y - \frac{\mathrm{i}p_0\lambda_0^2}{32\pi^2}\alpha(z)\alpha'(z)x + \frac{\mathrm{i}p_0\lambda_0^2}{16\pi^2}\alpha''(z)y\right. \\
&\quad -\frac{\mathrm{i}\lambda_0^2}{256\pi^2}\alpha'''(z)p_y r_\perp^2 - \frac{\mathrm{i}\lambda_0^2}{256\pi^2}\alpha'''(z)y\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp\right. \\
&\quad \left.+\boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right) - \frac{\mathrm{i}p_0\lambda_0^2}{128\pi^2}\alpha^{(4)}(z)yr_\perp^2 \\
&\quad \left.-\frac{p_0\lambda_0^2}{256\pi^2}\left(3\alpha'(z)\alpha''(z) + \mathrm{i}\alpha(z)\alpha'''(z)\right)xr_\perp^2\right\}\Sigma_y \\
&\quad +\left\{-\frac{p_0\lambda_0}{2\pi}\alpha(z) - \frac{\lambda_0}{4\pi p_0}\alpha(z)\hat{p}_\perp^2\right. \\
&\quad +\left(\frac{p_0\lambda_0}{8\pi}\left(\alpha''(z) - 2\alpha(z)^3\right) - \frac{\mathrm{i}p_0\lambda_0^2}{128\pi^2}\alpha'''(z)\right)r_\perp^2 \\
&\quad +\frac{\lambda_0}{8\pi}\alpha'(z)\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp + \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right) + \frac{\lambda_0}{2\pi}\alpha(z)^2\hat{L}_z \\
&\quad +\frac{\lambda_0}{32\pi p_0}\alpha''(z)\left(\hat{p}_\perp^2 r_\perp^2 + r_\perp^2\hat{p}_\perp^2\right)
\end{aligned}
$$

$$
\begin{aligned}
& -\frac{3\lambda_0}{16\pi}\alpha(z)\alpha''(z)\hat{L}_z r_\perp^2 + \frac{3p_0\lambda_0}{32\pi}\alpha(z)^2\alpha''(z)r_\perp^4 \\
& -\frac{\lambda_0}{128\pi}\alpha'''(z)\left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp + \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)r_\perp^2 \\
& -\frac{\lambda_0}{128\pi}\alpha'''(z)\left(\left\{p_x\,, xr_\perp^2\right\} + \left\{p_y\,, yr_\perp^2\right\}\right)\Big\}\,\Sigma_z \\
& -\frac{\mathrm{i}\lambda_0}{48\pi p_0}\alpha'(z)\chi\beta\left\{\boldsymbol{\alpha}_\perp\cdot\boldsymbol{p}_\perp, \left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp + \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)\right\} \\
& -\frac{\mathrm{i}\lambda_0}{32\pi}\alpha(z)\alpha'(z)\chi\beta\left\{\boldsymbol{\alpha}_\perp\cdot\boldsymbol{r}_\perp\,, \left(\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp + \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)\right\} \\
& +\frac{\mathrm{i}p_0\lambda_0}{4\pi}\alpha'(z)\chi\beta\left(\alpha_x y - \alpha_y x\right) \\
& +\left(-\frac{\mathrm{i}p_0\lambda_0}{32\pi}\alpha'''(z) + \frac{\mathrm{i}p_0\lambda_0^2}{64\pi^2}\alpha(z)\alpha''(z)\right) \\
& \quad\times\chi\beta\left(\alpha_x y - \alpha_y x\right)r_\perp^2 \\
& +\frac{p_0\lambda_0^2}{64\pi^2}\chi\beta\alpha_z\left(2\alpha'(z)^2 r_\perp^2 - \frac{1}{2}\alpha'(z)\alpha'''(z)r_\perp^4\right) \\
& +\frac{\lambda_0^2}{256\pi^2}\alpha'''(z)\chi\beta\left[\alpha_x\hat{p}_y - \alpha_y\hat{p}_x\,, r_\perp^2\right] \\
& \eta = \frac{E + m_0c^2}{cp_0}, \qquad \alpha(z) = \frac{qB(z)}{2p_0}\,.
\end{aligned}
\tag{3.69}
$$

Comparing with the scalar case it is seen that the difference in the scalar part ($\sim I$) lies only in the λ_0-dependent term. Thus as already noted, even the scalar approximation of the Dirac theory is, in principle, different from the Klein-Gordon theory, though it is only a slight difference exhibited in the λ_0-dependent terms. The matrix part in $\hat{\mathcal{H}}_o$ in the Dirac theory, $\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)}$, adds to the deviation from the Klein-Gordon theory. The computation of aberrations is exactly like in the scalar case using the interaction picture. Without further ado, let us just note that the position aberration $\left(\delta\boldsymbol{r}_\perp\right)_{(3)}(z_0)$ gets additional contributions of every type from the matrix part $\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)}$. For example, the additional spherical-aberration-type contribution is:

$$
\begin{aligned}
& C_s^{(\lambda_0,\sigma)}\langle\hat{\boldsymbol{p}}_\perp\hat{p}_\perp^2\Sigma_z\rangle(z_0)/p_0^3 \\
& = \left(C_s^{(\lambda_0,\sigma)}/p_0^3\right)\Big\{\left\langle\psi_1(z_o)\left|\hat{\boldsymbol{p}}_\perp\hat{p}_\perp^2\right|\psi_1(z_o)\right\rangle \\
& \quad - \left\langle\psi_2(z_o)\left|\hat{\boldsymbol{p}}_\perp\hat{p}_\perp^2\right|\psi_2(z_o)\right\rangle + \left\langle\psi_3(z_o)\left|\hat{\boldsymbol{p}}_\perp\hat{p}_\perp^2\right|\psi_3(z_o)\right\rangle
\end{aligned}
$$

$$
\begin{aligned}
& -\left\langle \psi_4(z_0)\left|\hat{p}_\perp \hat{p}_\perp^2\right| \psi_4(z_0)\right\rangle\Big\} \\
C_s^{(\lambda_0,\sigma)} &= \frac{\lambda_0}{16\pi}\int_{z_0}^{z_i} dz\,\left\{6\alpha^2\alpha'' h^4 - \alpha''' h^3 h' + 4\alpha'' h^2 h'^2\right\}\,,
\end{aligned}
\tag{3.70}
$$

where h is the 'classical' $h_p(z, z_0)$. Obviously, such a contribution, with unequal weights for the four spinor components, would depend on the nature of $|\psi(z_0)\rangle$ with respect to spin.

3.2.3 Magnetic Quadrupole Lens

Now, for the ideal magnetic quadrupole lens,

$$
\hat{\mathcal{H}}_o = -p_0 + \hat{\mathbf{H}}_{o,p} + \hat{\mathbf{H}}_{o,a} + \hat{\mathcal{H}}_o^{(\lambda_0)} + \hat{\mathcal{H}}_o^{(\lambda_0,\sigma)} \tag{3.71}
$$

$$
\hat{\mathbf{H}}_{o,p} = \frac{1}{2p_0}\hat{p}_\perp^2 - \frac{1}{2}qQ_m\left(x^2 - y^2\right) \tag{3.72}
$$

$$
\hat{\mathbf{H}}_{o,a} \approx \frac{1}{8p_0^3}\hat{p}_\perp^4 \tag{3.73}
$$

$$
\hat{\mathcal{H}}_o^{(\lambda_0)} \approx \frac{\lambda_0^2 q Q_m^2}{32\pi^2 p_0} r_\perp^2 \tag{3.74}
$$

$$
\begin{aligned}
\hat{\mathcal{H}}_o^{(\lambda_0,\sigma)} \approx\; & \frac{\lambda_0 q Q_m}{4\pi p_0}\left(x\hat{p}_y + y\hat{p}_x\right)\Sigma_z + \frac{\mathrm{i}\lambda_0^2 q Q_m}{4\pi^2 p_0}\left(\hat{p}_y\Sigma_x + \hat{p}_x\Sigma_y\right) \\
& + \frac{\mathrm{i}\lambda_0 q Q_m}{4\pi}\beta\chi\left(x\alpha_x - y\alpha_y\right) \\
& + \frac{\mathrm{i}\lambda_0 q Q_m}{4\pi p_0^2}\beta\chi\left\{\left(\hat{p}_y\alpha_x - \hat{p}_x\alpha_y\right), \left(x\hat{p}_y + y\hat{p}_x\right)\right\}\,.
\end{aligned}
\tag{3.75}
$$

Again, it is seen that, the λ_0-dependent scalar term, $\hat{\mathcal{H}}_o^{(\lambda_0)}$ is different from the corresponding one in the Klein-Gordon theory.

Chapter 4

Spin dynamics of the Dirac-particle beam

4.1 Introduction

In the previous chapters we developed the formalism for the quantum theory of charged-particle beam optics with applications mainly to the problem of imaging. In this chapter we shall be more interested in the applications of the theory to accelerator optics, particularly polarized beams and taking into account the anomalous magnetic moment of the Dirac particle. Here we are concerned only with the changes in the average values of the observables like, position, momenta and spin along the optic axis. To this end, it is best to work with the two-component formalism which is obtained as a result of the Foldy-Wouthuysen-like transformations without going back to the canonical Dirac representation as was done in the previous chapter to enable the interpretation of the imaging in terms of the familiar four-component wavefunction. Further, to study the spin dynamics of the beam particle one has to define the spin in the rest frame of the particle as is usually done in accelerator physics. It may be noted that in electron optical imaging we are not bothered about the spin of the particle; what we studied in the last chapter was the effect of the spin on the optical characteristics of the system. In accelerator physics one is interested in the actual value of the spin so that the polarization of the beam can be maintained at the desired level. The formalism of the present chapter, slightly different from that of the previous chapter, for the reasons just mentioned, is developed to suit the accelerator beam optics. This chapter will be almost self-consistent.

In accelerator physics it is customary to treat the beam optics essentially in terms of classical mechanics. As is well-known, the main framework for studying the spin dynamics and beam polarization in accelerator physics is essentially based on the well-known quasiclassical Thomas-Bargmann-Michel-Telegdi (Thomas-BMT) equation ([61, 62])(see, *e.g.*, [22] for a review). The other aspects, such as the quantum fluctuations of the trajectory and the radiative polarization have been approached using the quantized nature of radiation and solutions of the Dirac equation (see, *e.g.*, [63]-[67] and references therein). There have been several different approaches, independent of the beam optics, to understand the Thomas-BMT equation based on the Dirac equation (see, *e.g.*, [63, 68, 69] and references therein). Derbenev and Kondratenko [70] derived a semiclassical theory to describe, in a unified way, the orbital dynamics, polarization, and radiative processes, of a Dirac particle with anomalous magnetic moment starting with the Dirac equation and employing the Foldy-Wouthuysen technique (see also [67] where the same result is achieved using the Pauli reduction of the Dirac equation; see [71] for a discussion of the Derbenev-Kondratenko Hamiltonian within the context of classical theory of relativistic spin-orbit systems). The Derbenev-Kondratenko formalism has been the starting point for the development of a completely classical treatment of beam optics including spin with spin components taken as classical variables (see [72, 73]). In the recent work of Conte *et al.* [23] a fully quantum mechanical understanding of the *beam optics* has been achieved based on the Dirac equation by casting the Dirac equation directly into an accelerator optical form with the aid of the Foldy-Wouthuysen-like technique employed in the previous chapters.

Starting with the standard Dirac-Pauli equation for a spin-$\frac{1}{2}$ particle with anomalous magnetic moment it is possible to obtain a representation in which the effective 'accelerator optical' Hamiltonian accounts, in a unified way, for both the orbital (the Lorentz and the Stern-Gerlach forces) and the spin (the Thomas-BMT equation) motions. The general theory, developed for any magnetic element with straight optic axis and up to the lowest order (paraxial approximation), is illustrated by computing the transfer maps for phase-space and spin components in the cases normal magnetic quadrupole and skew magnetic quadrupole lenses. The quantum mechanics of Stern-Gerlach kicks is also discussed. The formalism treats the beam optics at the level of single particle dynamics, considers the electromagnetic field as classical and disregards the radiation aspects.

4.2 Formalism

We are interested in studying the spin dynamics and optics of a monoenergetic paraxial Dirac-particle beam transported through a magnetic optical element with straight axis comprising the static field $\boldsymbol{B} = \mathrm{curl}\, \boldsymbol{A}$ associated with a vector potential $\boldsymbol{A}$. Let us consider the Dirac particle to have mass m_0, charge q and anomalous magnetic moment μ_a. The beam propagation is governed by the stationary Dirac equation

$$\hat{\mathrm{H}}_D \left|\psi_D\right\rangle = E \left|\psi_D\right\rangle \,, \tag{4.1}$$

where $\left|\psi_D\right\rangle$ is the time-independent 4-component Dirac spinor, E is the energy of the beam particle and the Hamiltonian $\hat{\mathrm{H}}_D$, including the Pauli term is given by

$$\hat{\mathrm{H}}_D = \beta m_0 c^2 + c\boldsymbol{\alpha} \cdot \left(-i\hbar\boldsymbol{\nabla} - q\boldsymbol{A}\right) - \mu_a \beta \boldsymbol{\Sigma} \cdot \boldsymbol{B} \,,$$

$$\beta = \left(\begin{array}{cc} \mathbb{1} & \mathrm{O} \\ \mathrm{O} & -\mathbb{1} \end{array}\right) , \quad \boldsymbol{\alpha} = \left(\begin{array}{cc} \mathrm{O} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathrm{O} \end{array}\right) , \quad \boldsymbol{\Sigma} = \left(\begin{array}{cc} \boldsymbol{\sigma} & \mathrm{O} \\ \mathrm{O} & \boldsymbol{\sigma} \end{array}\right) ,$$

$$
\begin{aligned}
\mathbb{1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{o} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad I = \begin{pmatrix} \mathbb{1} & \mathbf{o} \\ \mathbf{o} & \mathbb{1} \end{pmatrix}, \\
\sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{aligned}
\tag{4.2}
$$

Note that we are dealing with the scattering states of the time-independent Hamiltonian $\hat{\mathrm{H}}_D$ with conserved positive energy

$$
E = +\sqrt{m_0^2 c^4 + c^2 p_0^2}, \quad p_0 = |\boldsymbol{p}_0|, \tag{4.3}
$$

where $\boldsymbol{p}_0$ is the momentum of the beam particle entering the system from the field-free input region. We shall consider the beam to be paraxial and moving along the positive z-direction such that for any constituent particle of the beam

$$
p_0 \approx p_z > 0, \quad |p_x| \ll p_0, \quad |p_y| \ll p_0. \tag{4.4}
$$

We shall use the right handed Cartesian coordinate system with z pointing along the design trajectory, y as the vertical coordinate and x as the horizontal transverse coordinate. Note that in accelerator physics there are different conventions used for the choice of the coordinate frame.

Since we want to know the changes in the beam parameters along the optic axis of the system (*i.e.*, the $+z$-direction) we have to study the Dirac equation (4.1) rewritten as

$$
\mathrm{i}\hbar \frac{\partial}{\partial z} |\psi_D\rangle = \hat{\mathcal{H}}_D |\psi_D\rangle , \tag{4.5}
$$

i.e., we want to know how the Dirac wavefunction satisfying (4.1) evolves with z. If we assume that for any constituent particle of the beam, scattered by the static field of the optical element, the probability of location at the transverse plane at z, namely $\int d^2\boldsymbol{r}_\perp \sum_{i=1}^4 |\psi_{Di}(\boldsymbol{r}_\perp, z)|^2$, is almost a constant in the region of interest, then, one can consider $(\psi_{D1}(\boldsymbol{r}_\perp, z), \psi_{D2}(\boldsymbol{r}_\perp, z), \psi_{D3}(\boldsymbol{r}_\perp, z), \psi_{D4}(\boldsymbol{r}_\perp, z))$, apart from a common normalization factor, as the components of a spinor wavefunction in the transverse plane at z, and regard z as a parameter evolving along the optic axis of the system. We multiply (4.1) from left by α_z/c and rearrange the terms to get the desired form (4.5) : The result is that

$$\begin{aligned}
\mathcal{H}_D &= -p_0\beta\chi\alpha_z - qA_z I + \alpha_z\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp + (\mu_a/c)\beta\alpha_z\boldsymbol{\Sigma}\cdot\boldsymbol{B}\,,\\
\chi &= \begin{pmatrix} \xi\mathbb{1} & \mathbf{o} \\ \mathbf{o} & -\xi^{-1}\mathbb{1}\end{pmatrix}, \quad \xi = \sqrt{\frac{E+m_0c^2}{E-m_0c^2}}\,,\\
\hat{\boldsymbol{\pi}}_\perp &= (-\mathrm{i}\hbar\boldsymbol{\nabla}_\perp - q\boldsymbol{A}_\perp) = (\hat{\boldsymbol{p}}_\perp - q\boldsymbol{A}_\perp)\,.
\end{aligned} \tag{4.6}$$

Next we make the transformation [8]:

$$|\psi_D\rangle \longrightarrow |\psi'\rangle = M\,|\psi_D\rangle\,. \tag{4.7}$$

Recollecting that,

$$M = \frac{1}{\sqrt{2}}(I+\chi\alpha_z)\,, \qquad M^{-1} = \frac{1}{\sqrt{2}}(I-\chi\alpha_z)\,, \tag{4.8}$$

one has

$$M(\beta\chi\alpha_z)M^{-1} = \beta\,. \tag{4.9}$$

This turns (4.5) into

$$\mathrm{i}\hbar\frac{\partial}{\partial z}\,|\psi'\rangle = \hat{\mathcal{H}}'\,|\psi'\rangle\,, \quad \hat{\mathcal{H}}' = M\hat{\mathcal{H}}_D M^{-1} = -p_0\beta + \hat{\mathcal{E}} + \hat{\mathcal{O}}\,, \tag{4.10}$$

with the matrix elements of $\hat{\mathcal{E}}$ and $\hat{\mathcal{O}}$ given by

$$\begin{aligned}
\hat{\mathcal{E}}_{11} &= -qA_z\mathbb{1} - (\mu_a/2c)\left\{\left(\xi+\xi^{-1}\right)\boldsymbol{\sigma}_\perp\cdot\boldsymbol{B}_\perp + \left(\xi-\xi^{-1}\right)\sigma_z B_z\right\}\,,\\
\hat{\mathcal{E}}_{12} &= \hat{\mathcal{E}}_{21} = \mathbf{o}\,,\\
\hat{\mathcal{E}}_{22} &= -qA_z\mathbb{1} - (\mu_a/2c)\left\{\left(\xi+\xi^{-1}\right)\boldsymbol{\sigma}_\perp\cdot\boldsymbol{B}_\perp - \left(\xi-\xi^{-1}\right)\sigma_z B_z\right\}\,,
\end{aligned} \tag{4.11}$$

and

$$\hat{\mathcal{O}}_{11} = \hat{\mathcal{O}}_{22} = \mathbf{o}\,,$$

$$
\begin{aligned}
\hat{\mathcal{O}}_{12} &= \xi \left[\boldsymbol{\sigma}_{\perp} \cdot \hat{\boldsymbol{\pi}}_{\perp} - (\mu_a/2c)\left\{\mathrm{i}\left(\xi - \xi^{-1}\right)(B_x\sigma_y - B_y\sigma_x)\right.\right. \\
&\qquad\qquad \left.\left. - \left(\xi + \xi^{-1}\right) B_z \mathbb{1}\right\}\right] , \\
\hat{\mathcal{O}}_{21} &= -\xi^{-1} \left[\boldsymbol{\sigma}_{\perp} \cdot \hat{\boldsymbol{\pi}}_{\perp} + (\mu_a/2c)\left\{\mathrm{i}\left(\xi - \xi^{-1}\right)(B_x\sigma_y - B_y\sigma_x)\right.\right. \\
&\qquad\qquad \left.\left. + \left(\xi + \xi^{-1}\right) B_z \mathbb{1}\right\}\right] . \qquad (4.12)
\end{aligned}
$$

Recall that the transformation (4.7) is such that for a paraxial Dirac spinor propagating in the $+z$-direction $|\psi'\rangle$ is such that its lower spinor components are very small compared to the upper spinor components. To see this, let us consider the standard free Dirac plane-wave associated with positive energy E, namely,

$$
\begin{aligned}
\begin{pmatrix} \psi_{FD1}(\boldsymbol{r}_\perp, z) \\ \psi_{FD2}(\boldsymbol{r}_\perp, z) \\ \psi_{FD3}(\boldsymbol{r}_\perp, z) \\ \psi_{FD4}(\boldsymbol{r}_\perp, z) \end{pmatrix} &= \frac{1}{4}\sqrt{\frac{\xi c p_0}{\pi^3 \hbar^3 E}} \begin{pmatrix} s_+ \\ s_- \\ \{s_- p_- + s_+ p_z\}/\xi p_0 \\ \{s_+ p_+ - s_- p_z\}/\xi p_0 \end{pmatrix} \\
&\qquad \times \exp\left\{\frac{\mathrm{i}}{\hbar}\left(\boldsymbol{p}_\perp \cdot \boldsymbol{r}_\perp + p_z z\right)\right\} , \\
\boldsymbol{r}_\perp &= (x, y) , \quad |s_+|^2 + |s_-|^2 = 1 , \\
p_+ &= p_x + \mathrm{i} p_y , \quad p_- = p_x - \mathrm{i} p_y . \qquad (4.13)
\end{aligned}
$$

Correspondingly,

$$
\begin{aligned}
\begin{pmatrix} \psi'_{F1}(\boldsymbol{r}_\perp, z) \\ \psi'_{F2}(\boldsymbol{r}_\perp, z) \\ \psi'_{F3}(\boldsymbol{r}_\perp, z) \\ \psi'_{F4}(\boldsymbol{r}_\perp, z) \end{pmatrix} &= \frac{1}{4}\sqrt{\frac{\xi c p_0}{2\pi^3 \hbar^3 E}} \begin{pmatrix} \{s_+(p + p_z) + s_- p_-\}/p_0 \\ \{s_-(p + p_z) - s_+ p_+\}/p_0 \\ -\{s_+(p - p_z) - s_- p_-\}/\xi p_0 \\ \{s_-(p - p_z) + s_+ p_+\}/\xi p_0 \end{pmatrix} \\
&\qquad \times \exp\left\{\frac{\mathrm{i}}{\hbar}\left(\boldsymbol{p}_\perp \cdot \boldsymbol{r}_\perp + p_z z\right)\right\} , \qquad (4.14)
\end{aligned}
$$

and for a paraxial plane-wave moving in the positive z-direction, satisfying the condition (4.4), the upper spinor components of $|\psi'\rangle_F$, namely, $|\psi'_1\rangle_F$ and $|\psi'_2\rangle_F$, are obviously very large compared to its lower spinor components $|\psi'_3\rangle_F$ and $|\psi'_4\rangle_F$.

Like before, we shall follow the Foldy-Wouthuysen-like transformation technique to expand the beam-optical Hamiltonian in (4.10) into paraxial and aberrating parts.

To this end, we substitute in (4.10)

$$\left|\psi'\right\rangle = \exp\left(\frac{1}{2p_0}\beta\hat{\mathcal{O}}\right)\left|\psi^{(1)}\right\rangle . \qquad (4.15)$$

The resulting equation for $\left|\psi^{(1)}\right\rangle$ is

$$\begin{aligned}
\mathrm{i}\hbar\frac{\partial}{\partial z}\left|\psi^{(1)}\right\rangle &= \hat{\mathcal{H}}^{(1)}\left|\psi^{(1)}\right\rangle , \\
\hat{\mathcal{H}}^{(1)} &= \exp\left(-\frac{1}{2p_0}\beta\hat{\mathcal{O}}\right)\hat{\mathcal{H}}'\exp\left(\frac{1}{2p_0}\beta\hat{\mathcal{O}}\right) \\
&\quad -\mathrm{i}\hbar\exp\left(-\frac{1}{2p_0}\beta\hat{\mathcal{O}}\right)\frac{\partial}{\partial z}\left\{\exp\left(\frac{1}{2p_0}\beta\hat{\mathcal{O}}\right)\right\} \\
&= -p_0\beta + \hat{\mathcal{E}}^{(1)} + \hat{\mathcal{O}}^{(1)} , \\
\hat{\mathcal{E}}^{(1)} &= \hat{\mathcal{E}} - \frac{1}{2p_0}\beta\hat{\mathcal{O}}^2 + \cdots , \\
\hat{\mathcal{O}}^{(1)} &= -\frac{1}{2p_0}\beta\left\{\left[\hat{\mathcal{O}},\hat{\mathcal{E}}\right] + \mathrm{i}\hbar\frac{\partial}{\partial z}\hat{\mathcal{O}}\right\} + \cdots .
\end{aligned} \qquad (4.16)$$

As stated earlier we shall confine our study to the paraxial case. So we stop with the above first step which corresponds to the paraxial approximation. Let us write down explicitly, for later use, the 11-block element of $\hat{\mathcal{H}}^{(1)}$:

$$\begin{aligned}
\hat{\mathcal{H}}_{11}^{(1)} &= -p_0\mathbb{1} + \hat{\mathcal{E}}_{11}^{(1)} \\
&= \left\{-p_0 - qA_z + \frac{1}{2p_0}\hat{\pi}_\perp^2 - \frac{\epsilon\hbar^2}{4p_0^2}(\mathrm{curl}\,\boldsymbol{B})_z + \frac{\epsilon^2\hbar^2}{8p_0^3}\left(B_\perp^2 + \gamma^2 B_z^2\right)\right\}\mathbb{1} \\
&\quad -\frac{1}{p_0}\left\{(q+\epsilon)B_z S_z + \gamma\epsilon\boldsymbol{B}_\perp\cdot\boldsymbol{S}_\perp\right\} \\
&\quad +\frac{\epsilon}{2p_0^2}\left\{\gamma\left(B_z\boldsymbol{S}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp + \boldsymbol{S}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp B_z\right) - \left(\boldsymbol{B}_\perp\cdot\hat{\boldsymbol{\pi}}_\perp + \hat{\boldsymbol{\pi}}_\perp\cdot\boldsymbol{B}_\perp\right)S_z\right\} ,
\end{aligned}$$

$$\hat{\pi}_\perp^2 = \hat{\pi}_x^2 + \hat{\pi}_y^2 , \quad \epsilon = 2m_0\mu_a/\hbar , \quad \gamma = E/m_0c^2 , \quad \boldsymbol{S} = \frac{1}{2}\hbar\boldsymbol{\sigma} . \qquad (4.17)$$

Before proceeding further, let us find out the nature of $\left|\psi^{(1)}\right\rangle$ by looking at the field-free case again. For $\left|\psi\right\rangle_F$ in (4.13)

$$\left|\psi^{(1)}\right\rangle_F = \exp\left(-\frac{1}{2p_0}\beta\chi\boldsymbol{\alpha}_\perp\cdot\hat{\boldsymbol{p}}_\perp\right)\left|\psi'\right\rangle_F$$

$$\approx \left(I - \frac{1}{2p_0}\beta\chi\boldsymbol{\alpha}_\perp \cdot \hat{\boldsymbol{p}}_\perp\right)|\psi'\rangle_F ,$$

$$\begin{pmatrix} \psi_{F1}^{(1)}(\boldsymbol{r}_\perp, z) \\ \psi_{F2}^{(1)}(\boldsymbol{r}_\perp, z) \\ \psi_{F3}^{(1)}(\boldsymbol{r}_\perp, z) \\ \psi_{F4}^{(1)}(\boldsymbol{r}_\perp, z) \end{pmatrix} \approx \frac{1}{4}\sqrt{\frac{\xi c p_0}{2\pi^3\hbar^3 E}}$$
$$\times \begin{pmatrix} s_+\left\{1 + \frac{p_z}{p_0} - \frac{p_\perp^2}{2p_0^2}\right\} + \frac{1}{2}s_-\left\{\left(1+\frac{p_z}{p_0}\right)\frac{p_-}{p_0}\right\} \\ s_-\left\{1 + \frac{p_z}{p_0} + \frac{p_\perp^2}{2p_0^2}\right\} - \frac{1}{2}s_+\left\{\left(1+\frac{p_z}{p_0}\right)\frac{p_+}{p_0}\right\} \\ -\frac{1}{\xi}\left[s_+\left\{1 - \frac{p_z}{p_0} - \frac{p_\perp^2}{2p_0^2}\right\} + \frac{1}{2}s_-\left\{\left(1-\frac{p_z}{p_0}\right)\frac{p_-}{p_0}\right\}\right] \\ \frac{1}{\xi}\left[s_-\left\{1 - \frac{p_z}{p_0} - \frac{p_\perp^2}{2p_0^2}\right\} + \frac{1}{2}s_+\left\{\left(1-\frac{p_z}{p_0}\right)\frac{p_+}{p_0}\right\}\right] \end{pmatrix}$$
$$\times \exp\left\{\frac{i}{\hbar}\left(\boldsymbol{p}_\perp \cdot \boldsymbol{r}_\perp + p_z z\right)\right\} , \tag{4.18}$$

showing clearly that the transformation (4.15) keeps the upper spinor components of $\left|\psi^{(1)}\right\rangle$ large compared to its lower spinor components.

Since the lower pair of components of $\left|\psi^{(1)}\right\rangle$ $\left(\left|\psi_3^{(1)}\right\rangle\right.$ and $\left.\left|\psi_4^{(1)}\right\rangle\right)$ are almost vanishing compared to the upper pair $\left(\left|\psi_1^{(1)}\right\rangle\right.$ and $\left.\left|\psi_2^{(1)}\right\rangle\right)$ and the odd part of $\hat{\mathcal{H}}^{(1)}$ is negligible compared to its even part we can effectively introduce a Pauli-like two-component spinor formalism based on the representation (4.16). Naming the two-component spinor comprising the upper pair of components of $\left|\psi^{(1)}\right\rangle$ as $\left|\tilde{\psi}\right\rangle$ and calling $\hat{\mathcal{H}}_{11}^{(1)}$ as $\hat{\tilde{\mathcal{H}}}$ it is clear from (4.16) and (4.17) that we can write

$$i\hbar\frac{\partial}{\partial z}\left|\tilde{\psi}\right\rangle = \hat{\tilde{\mathcal{H}}}\left|\tilde{\psi}\right\rangle , \qquad \left|\tilde{\psi}\right\rangle = \begin{pmatrix} \left|\tilde{\psi}_1\right\rangle \\ \left|\tilde{\psi}_2\right\rangle \end{pmatrix} ,$$
$$\hat{\tilde{\mathcal{H}}} \approx \left(-p_0 - qA_z + \frac{1}{2p_0}\hat{\pi}_\perp^2\right) - \frac{1}{p_0}\left\{(q+\epsilon)B_zS_z + \gamma\epsilon\boldsymbol{B}_\perp \cdot \boldsymbol{S}_\perp\right\} , \tag{4.19}$$

where $\hat{\tilde{H}}$ has been approximated by keeping only terms up to first order in $1/p_0$ (see (4.17)), consistent with the assumption of paraxiality condition (4.4) for the beam. Throughout this chapter we shall approximate the various expressions by keeping only up to the lowest order nontrivial terms consistent with the paraxiality condition for the beam and the approximation symbol ($\approx$) will usually imply this, wherever relevant, even if not stated explicitly.

Up to now, all the observables, the field components, time, etc., are defined with reference to the laboratory frame. But, as is well-known, in the covariant description the spin of the Dirac particle has simple operator representation in terms of the Pauli matrices only in a frame at which the particle is at rest. So, as is usual, we shall prefer to define spin with reference to the instantaneous rest frame of the particle while keeping the other observables, field components, time, etc., defined with reference to the laboratory frame. To this end, we transform the two-component $\left|\tilde{\psi}\right\rangle$ to an 'accelerator optics representation' $\left|\psi^{(A)}\right\rangle$ defined by

$$\left|\tilde{\psi}\right\rangle = \exp\left\{\frac{\mathrm{i}}{2p_0}\left(\hat{\pi}_x\sigma_y - \hat{\pi}_y\sigma_x\right)\right\}\left|\psi^{(A)}\right\rangle . \tag{4.20}$$

The reason for the choice of this transformation will become clear shortly. Now, the z-evolution equation for $|\psi^{(A)}\rangle$ is

$$\begin{aligned}
\mathrm{i}\hbar\frac{\partial}{\partial z}\left|\psi^{(A)}\right\rangle &= \hat{H}^{(A)}\left|\psi^{(A)}\right\rangle , \\
\hat{H}^{(A)} &\approx \left(-p_0 - qA_z + \frac{1}{2p_0}\hat{\pi}_\perp^2\right) + \frac{\gamma m_0}{p_0}\underline{\boldsymbol{\Omega}}_s \cdot \boldsymbol{S} , \\
\text{with } \underline{\boldsymbol{\Omega}}_s &= -\frac{1}{\gamma m_0}\left\{q\boldsymbol{B} + \epsilon\left(\boldsymbol{B}_\parallel + \gamma\boldsymbol{B}_\perp\right)\right\} ,
\end{aligned} \tag{4.21}$$

where $\boldsymbol{B}_\parallel$ and $\boldsymbol{B}_\perp$ are the components of $\boldsymbol{B}$ along the z-axis and perpendicular to it. When $q = \pm e$ we can write $\epsilon = qa = q(g-2)/2$ where g and a are, respectively, the gyromagnetic ratio and the magnetic anomaly of the particle; for the neutron $\epsilon = g|e|/2$. It may be noted that the accelerator optical quantum Hamiltonian $\hat{H}^{(A)}$ is hermitian though $\tilde{\mathcal{H}}_D$ in (4.6) is nonhermitian. The nonunitary similarity transformations we have made have resulted in this change and the hermiticity of $\hat{H}^{(A)}$ implies the approximate constancy of the total intensity of the beam in any transverse plane along the optic axis. It should be realized that the spin part of $\hat{H}^{(A)}$ corresponds to the beam optical and paraxial version of the Thomas-BMT spin Hamiltonian : note that $\boldsymbol{B}_{\parallel}$ and $\boldsymbol{B}_{\perp}$ in the usual Thomas-BMT vector $\boldsymbol{\Omega}_s$ refer to the parallel and the perpendicular components with respect to the instantaneous velocity of the particle whereas in $\underline{\boldsymbol{\Omega}}_s$ in (4.21) they refer to the parallel and perpendicular components with respect to the predominant direction of propagation of the particle. The Thomas-BMT part of $\hat{H}^{(A)}$ is also valid up to first order in $\hbar$. To get higher order corrections, in terms of $\hat{\pi}_{\perp}/p_0$ and $\hbar$, we have to go beyond the first Foldy-Wouthuysen-like transformation (4.15). It may also be noted that $\hat{H}^{(A)}$ is the accelerator optical version of the Derbenev-Kondratenko Hamiltonian [70] for the Dirac particle, under the paraxial approximation.

Since the z-evolution of $\left|\psi^{(A)}\right\rangle$ is unitary we can associate the beam with a wavefunction normalized in such a way that, at any z,

$$\left\langle \psi^{(A)}(z) \middle| \psi^{(A)}(z) \right\rangle = \sum_{i=1}^{2} \int d^2\boldsymbol{r}_{\perp} \left|\psi_i^{(A)}(\boldsymbol{r}_{\perp}, z)\right|^2 = 1 \,. \tag{4.22}$$

When the beam is described by a 2×2 statistical (density) matrix

$$\rho^{(A)} = \left(\begin{array}{cc} \rho_{11}^{(A)} & \rho_{12}^{(A)} \\ \rho_{21}^{(A)} & \rho_{22}^{(A)} \end{array} \right) , \tag{4.23}$$

with the normalization

$$\mathrm{Tr}\left(\rho^{(A)}(z)\right) = \sum_{i=1}^{2} \int d^2\boldsymbol{r}_{\perp} \left\langle \boldsymbol{r}_{\perp} \left| \rho_{ii}^{(A)}(z) \right| \boldsymbol{r}_{\perp} \right\rangle = 1 \,, \tag{4.24}$$

at any z, the accelerator optical z-evolution equation is

$$\mathrm{i}\hbar \frac{\partial}{\partial z} \rho^{(A)} = \left[\hat{H}^{(A)} , \rho^{(A)} \right] . \tag{4.25}$$

If the beam can be described as a pure state we would have $\rho^{(A)} = \left|\psi^{(A)}\right\rangle \left\langle \psi^{(A)}\right|$.

Let us now define the average of any observable O at the transverse plane at z to be given by

$$\left\langle \hat{O}^{(A)} \right\rangle (z) \;\; = \;\; \mathrm{Tr}\left(\rho^{(A)}(z) \hat{O}^{(A)}\right)$$

$$= \sum_{i,j=1}^{2} \int\int d^2\boldsymbol{r}_\perp d^2\boldsymbol{r}'_\perp \left\langle \boldsymbol{r}_\perp \left| \rho_{ij}^{(A)}(z) \right| \boldsymbol{r}'_\perp \right\rangle \left\langle \boldsymbol{r}'_\perp \left| \hat{O}_{ji}^{(A)} \right| \boldsymbol{r}_\perp \right\rangle , \tag{4.26}$$

where $\hat{O}^{(A)}$ is the operator representing O in the accelerator optical representation.

For any observable O, associated with the operator $\hat{O}_D$ in the standard Dirac representation (4.1) the corresponding $\hat{O}^{(A)}$ can be obtained as follows :

$$\begin{aligned} \hat{O}^{(A)} \ = \ & \text{the hermitian part of the 11} - \text{block element of} \\ & \left(\exp\left\{ -\frac{\mathrm{i}}{2p_0} \left(\hat{\pi}_x \Sigma_y - \hat{\pi}_y \Sigma_x \right) \right\} \right. \\ & \times \exp\left(-\frac{1}{2p_0} \beta \hat{\mathcal{O}} \right) M \hat{O}_D M^{-1} \exp\left(\frac{1}{2p_0} \beta \hat{\mathcal{O}} \right) \\ & \left. \times \exp\left\{ \frac{\mathrm{i}}{2p_0} \left(\hat{\pi}_x \Sigma_y - \hat{\pi}_y \Sigma_x \right) \right\} \right) . \end{aligned} \tag{4.27}$$

In the Dirac representation the operator for the spin unit vector corresponding to the spin as defined in the instantaneous rest frame of the particle (see [63]) is given by

$$\boldsymbol{S}_R = \frac{\hbar}{2} \begin{pmatrix} \boldsymbol{\sigma} - \frac{c^2(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\pi}} + \hat{\boldsymbol{\pi}}\cdot\boldsymbol{\sigma})}{2E(E+m_0c^2)} & \frac{c\hat{\boldsymbol{\pi}}}{E} \\ \frac{c\hat{\boldsymbol{\pi}}}{E} & -\boldsymbol{\sigma} + \frac{c^2(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\pi}} + \hat{\boldsymbol{\pi}}\cdot\boldsymbol{\sigma})}{2E(E+m_0c^2)} \end{pmatrix} . \tag{4.28}$$

If we now compute the corresponding operator $\boldsymbol{S}_R^{(A)}$ in the accelerator optical representation, using the formula (4.27), it is found that up to first order (paraxial) approximation

$$\boldsymbol{S}_R^{(A)} \approx \frac{\hbar}{2} \boldsymbol{\sigma} \tag{4.29}$$

as is desired. In the Dirac representation the position operator in free space can be taken to be given by the mean position operator as indicated by the Foldy-Wouthuysen-theory (or what is same as the Newton-Wigner position operator). In presence of the magnetic field we can extend this position operator by the replacement $\hat{\boldsymbol{p}} \longrightarrow \hat{\boldsymbol{\pi}}$ and symmetrization (to make it hermitian). Then, the operator for the transverse position coordinate in the accelerator optical representation becomes just the canonical position operator $\boldsymbol{r}_\perp$ in the first order approximation. From these considerations it is clear that in the accelerator optical evolution equation (4.21) $\boldsymbol{S}$ represents the spin as defined in the instantaneous rest frame of the particle; the field components and other operators are all defined with respect to the laboratory frame. It should be noted that in this formalism, with z as the evolution parameter (analogous to time t), $-\hat{H}^{(A)}$ corresponding to $-\mathrm{i}\hbar\frac{\partial}{\partial z}$, will represent $\hat{p}_z$, the z-component of canonical momentum operator (analogous to the energy operator); hence, the operator $-(\hat{H}^{(A)} + qA_z)$ will represent π_z the z-component of the kinetic momentum.

If we now work out the equations of motion for the average values of $\boldsymbol{r}_\perp$ using (4.25), they have to be consistent, *à la* Ehrenfest, with the traditional transfer map for the phase-space, including the transverse Stern-Gerlach kicks (see, *e.g.*, [25], [74]-[80]), in the paraxial approximation. The transfer map for the averages of spin components, in the lowest order approximation, has to be consistent with the Thomas-BMT equation. This is confirmed easily by a preliminary analysis as follows. From (4.25) and (4.26) we have, in general,

$$\frac{d}{dz}\left\langle\hat{O}^{(A)}\right\rangle(z) = -\frac{\mathrm{i}}{\hbar}\left\langle\left[\hat{O}^{(A)}\,,\,\hat{H}^{(A)}\right]\right\rangle(z) + \left\langle\frac{\partial}{\partial z}\hat{O}^{(A)}\right\rangle(z)\,. \tag{4.30}$$

To compare (4.30) with the time evolution of classical O we can use the correspondence

$$\frac{d}{dt}O \longrightarrow \frac{d}{dt}\left\langle\hat{O}^{(A)}\right\rangle \approx v_z\frac{d}{dz}\left\langle\hat{O}^{(A)}\right\rangle \approx \frac{p_0}{\gamma m_0}\frac{d}{dz}\left\langle\hat{O}^{(A)}\right\rangle\,, \tag{4.31}$$

since

$$\begin{aligned} v_z &= \frac{1}{\gamma m_0}\langle\hat{\pi}_z\rangle = -\frac{1}{\gamma m_0}\left\langle\hat{H}^{(A)} + qA_z\right\rangle \\ &= \frac{p_0}{\gamma m_0} - \frac{1}{2\gamma m_0 p_0}\left\langle\hat{\pi}_\perp^2\right\rangle - \frac{1}{p_0}\left\langle\underline{\Omega}_s\cdot\boldsymbol{S}\right\rangle \approx \frac{p_0}{\gamma m_0}\,. \end{aligned} \tag{4.32}$$

Then, for $\boldsymbol{r}_\perp$ we get

$$\frac{d}{dz}\langle\boldsymbol{r}_\perp\rangle \approx -\frac{\mathrm{i}}{\hbar}\left\langle\left[\boldsymbol{r}_\perp\,,\,\hat{H}^{(A)}\right]\right\rangle(z) = \frac{1}{p_0}\langle\hat{\boldsymbol{\pi}}_\perp\rangle\,, \tag{4.33}$$

and hence from (4.31),

$$\frac{d}{dt}\langle \boldsymbol{r}_\perp\rangle \approx -\frac{\mathrm{i}}{\hbar}\frac{p_0}{\gamma m_0}\left\langle\left[\boldsymbol{r}_\perp\,,\,\hat{H}^{(A)}\right]\right\rangle(z) = \frac{1}{\gamma m_0}\langle\hat{\boldsymbol{\pi}}_\perp\rangle\,, \tag{4.34}$$

identifying $\hat{\boldsymbol{\pi}}_\perp$ as the transverse kinetic momentum. From (4.33) it is clear that $\langle \boldsymbol{r}_\perp\rangle$ and $\langle\hat{\boldsymbol{\pi}}_\perp\rangle/p_0$ ($\langle\hat{\boldsymbol{p}}_\perp\rangle/p_0$ in the field-free regions) can be identified with the transverse position and slope of the classical ray corresponding to the wavepacket represented by $\rho^{(A)}$. For $\hat{\boldsymbol{\pi}}_\perp$, we have, with $\hat{\pi}_z \approx p_0$,

$$\begin{aligned}
\frac{d}{dz}\langle\hat{\boldsymbol{\pi}}_\perp\rangle &\approx -\frac{\mathrm{i}}{\hbar}\left\langle\left[\hat{\boldsymbol{\pi}}_\perp\,,\,\hat{H}^{(A)}\right]\right\rangle - q\left\langle\frac{\partial}{\partial z}\boldsymbol{A}_\perp\right\rangle \\
&\approx \frac{q}{p_0}\left\langle\frac{1}{2}\left(\hat{\boldsymbol{\pi}}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\boldsymbol{\pi}}\right)_\perp\right\rangle - \frac{\gamma m_0}{p_0}\left\langle\boldsymbol{\nabla}_\perp\left(\boldsymbol{\Omega}_s\cdot\boldsymbol{S}\right)\right\rangle \\
&= \frac{q}{p_0}\left\langle\frac{1}{2}\left(\hat{\boldsymbol{\pi}}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\boldsymbol{\pi}}\right)_\perp\right\rangle \\
&\qquad + \frac{1}{p_0}\left\langle\boldsymbol{\nabla}_\perp\left\{(q+\epsilon)B_zS_z + (q+\gamma\epsilon)\boldsymbol{B}_\perp\cdot\boldsymbol{S}_\perp\right\}\right\rangle\,,
\end{aligned} \tag{4.35}$$

and hence

$$\begin{aligned}
\frac{d}{dt}\langle\hat{\boldsymbol{\pi}}_\perp\rangle &= \frac{q}{\gamma m_0}\left\langle\frac{1}{2}\left(\hat{\boldsymbol{\pi}}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\boldsymbol{\pi}}\right)_\perp\right\rangle \\
&\qquad + \frac{1}{\gamma m_0}\left\langle\boldsymbol{\nabla}_\perp\left\{(q+\epsilon)B_zS_z + (q+\gamma\epsilon)\boldsymbol{B}_\perp\cdot\boldsymbol{S}_\perp\right\}\right\rangle\,.
\end{aligned} \tag{4.36}$$

Equation (4.36) is just in accordance with the quasiclassical equation for motion under the Lorentz and Stern-Gerlach forces up to the approximations considered. In the case of spin

$$\frac{d}{dz}\langle\boldsymbol{S}\rangle \approx -\frac{\mathrm{i}}{\hbar}\left\langle\left[\boldsymbol{S}\,,\,\hat{H}^{(A)}\right]\right\rangle = -\frac{\mathrm{i}}{\hbar}\frac{\gamma m_0}{p_0}\left\langle\left[\boldsymbol{S}\,,\,\boldsymbol{\Omega}_s\cdot\boldsymbol{S}\right]\right\rangle = \frac{\gamma m_0}{p_0}\left\langle\boldsymbol{\Omega}_s\times\boldsymbol{S}\right\rangle\,, \tag{4.37}$$

and thus,

$$\frac{d}{dt}\langle\boldsymbol{S}\rangle \approx -\frac{\mathrm{i}}{\hbar}\frac{p_0}{\gamma m_0}\left\langle\left[\boldsymbol{S}\,,\,\hat{H}^{(A)}\right]\right\rangle = -\frac{\mathrm{i}}{\hbar}\left\langle\left[\boldsymbol{S}\,,\,\boldsymbol{\Omega}_s\cdot\boldsymbol{S}\right]\right\rangle = \left\langle\boldsymbol{\Omega}_s\times\boldsymbol{S}\right\rangle\,, \tag{4.38}$$

as should be expected from the Thomas-BMT equation, of course up to the approximation we are concerned with. The vector $\boldsymbol{P}$ characterizing the polarization of the beam is given by the relation

$$\langle \boldsymbol{S} \rangle = \frac{\hbar}{2}\langle \boldsymbol{\sigma} \rangle = \frac{\hbar}{2}\boldsymbol{P} . \tag{4.39}$$

To obtain the required maps for transfer of the averages $(\langle \boldsymbol{r}_\perp \rangle, \langle \hat{\boldsymbol{\pi}}_\perp \rangle, \langle \boldsymbol{S} \rangle)$ across an optical element we can employ the quantum mechanical version [12] of the technique developed by Dragt *et al.* (see [15]-[21] and references therein) in the context of classical accelerator optics. We shall explain this in the following sections through the examples of quadrupolar magnetic lenses. Though we have taken $\langle \hat{\pi}_z \rangle \approx p$ in the above preliminary analysis, following (4.32), to understand the small variations in the longitudinal kinetic momentum, including the Stern-Gerlach kicks [25], a more careful analysis of the evolution of $\langle \hat{\pi}_z \rangle(z)$ along the z-axis is needed. We shall discuss this in the next section by examining the case of a general inhomogeneous magnetic field.

Before closing this section let us note that the Pauli-like two-component spinor formalism developed above is valid for all values of p_0, from the nonrelativistic to the extreme relativistic case; it becomes Pauli's two-component formalism in the nonrelativistic case when we can take $p_0 \approx \sqrt{2m_0(E - m_0c^2)}$.

4.3 Applications

4.3.1 Normal Magnetic Quadrupole

First, let us consider an ideal normal magnetic quadrupole lens field given by

$$\boldsymbol{B} = (-Gy, -Gx, 0), \tag{4.40}$$

associated with the vector potential

$$\boldsymbol{A} = \left(0, 0, \frac{1}{2}G\left(x^2 - y^2\right)\right) , \tag{4.41}$$

where G is assumed to be a constant in the lens region and zero outside. Let the z-coordinates of the xy-planes at the entrance and exit of the quadrupole magnet of length ℓ be $z_{\rm n}$ and $z_{\rm x}$ (the subscripts 'n' and 'x' denoting e'n'trance and e'x'it, respectively, and $\ell = z_{\rm x} - z_{\rm n}$). Throughout the present section we shall be working with the accelerator optical representation and shall omit the superscript (A).

Now, the basic accelerator optical Hamiltonian of the system is

$$\hat{H}(z) = \begin{cases} \hat{H}_F = -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2\,, & \text{for } z < z_{\rm n} \text{ and } z > z_{\rm x}\,, \\ \hat{H}_L(z) = -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2 - \frac{1}{2}qG\left(x^2 - y^2\right) + \frac{\eta p_0}{\ell}\left(y\sigma_x + x\sigma_y\right)\,, & \\ \qquad \text{for } z_{\rm n} \leq z \leq z_{\rm x}\,, \quad \text{with } \eta = (q + \gamma\epsilon)G\ell\hbar/2p_0^2\,. & \end{cases} \tag{4.42}$$

The subscripts F and L indicate, respectively, the field-free and the lens regions. Let us write $\hat{H}$ as a core part $\bar{\hat{H}}$ plus a perturbation part $\bar{\bar{\hat{H}}}$:

$$\hat{H}(z) = \bar{\hat{H}}(z) + \bar{\bar{\hat{H}}}(z)\,,$$

$$\bar{\hat{H}}(z) = \begin{cases} \bar{\hat{H}}_F \equiv \hat{H}_F\,, & \text{for } z < z_{\rm n} \text{ and } z > z_{\rm x}\,, \\ \bar{\hat{H}}_L(z) = -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2 - \frac{1}{2}qG\left(x^2 - y^2\right)\,, & \text{for } z_{\rm n} \leq z \leq z_{\rm x}\,. \end{cases}$$

$$\bar{\bar{\hat{H}}}(z) = \begin{cases} \bar{\bar{\hat{H}}}_F = 0\,, & \text{for } z < z_{\rm n} \text{ and } z > z_{\rm x}\,, \\ \bar{\bar{\hat{H}}}_L(z) = \frac{\eta p_0}{\ell}\left(y\sigma_x + x\sigma_y\right)\,, & \text{for } z_{\rm n} \leq z \leq z_{\rm x}\,. \end{cases} \tag{4.43}$$

A formal integration of the basic z-evolution equation (4.25) for ρ leads, in general, to

$$\rho(z) = \hat{U}\left(z, z_{\rm o}\right)\rho(z_{\rm o})\hat{U}^\dagger\left(z, z_{\rm o}\right)\,, \quad z \geq z_{\rm o}\,, \tag{4.44}$$

with the unitary z-propagator U given by

$$\hat{U}\left(z, z_{\rm o}\right) = \wp\left[\exp\left\{-\frac{\rm i}{\hbar}\int_{z_{\rm o}}^{z} d\zeta\,\hat{H}(\zeta)\right\}\right]\,, \tag{4.45}$$

where $\wp$ indicates the path-ordering of the exponential. Further, $\hat{U}$ is such that

$${\rm i}\hbar\frac{\partial}{\partial z}\hat{U}\left(z, z_{\rm o}\right) = \hat{H}(z)\hat{U}\left(z, z_{\rm o}\right)\,, \quad \hat{U}(z_{\rm o}, z_{\rm o}) = \hat{\mathcal{I}}\,, \tag{4.46}$$

where $\hat{\mathcal{I}}$ is the identity operator.

Let us now compute $\rho(z)$ *via* the interaction picture, familiar to us, from the earlier chapters. Defining

$$\rho_i(z) = \hat{\hat{U}}^{\dagger}(z, z_0)\,\rho(z)\hat{\hat{U}}(z, z_0)\,, \qquad \hat{\hat{U}}(z, z_0) = \wp\left[\exp\left\{-\frac{i}{\hbar}\int_{z_0}^{z} d\zeta\,\hat{\hat{H}}(\zeta)\right\}\right]\,, \tag{4.47}$$

we have

$$i\hbar\frac{\partial}{\partial z}\rho_i = \left[\hat{\hat{H}}_i\,, \rho_i\right], \qquad \hat{\hat{H}}_i = \hat{\hat{U}}^{\dagger}(z, z_0)\,\hat{\hat{H}}\hat{\hat{U}}(z, z_0)\,. \tag{4.48}$$

Then, since $\rho_i(z_0) = \rho(z_0)$,

$$\begin{aligned} \rho_i(z) &= \hat{\hat{U}}_i(z, z_0)\,\rho_i(z_0)\hat{\hat{U}}_i^{\dagger}(z, z_0) = \hat{\hat{U}}_i(z, z_0)\,\rho(z_0)\hat{\hat{U}}_i^{\dagger}(z, z_0)\,, \\ \hat{\hat{U}}_i(z, z_0) &= \wp\left[\exp\left\{-\frac{i}{\hbar}\int_{z_0}^{z} d\zeta\,\hat{\hat{H}}_i(\zeta)\right\}\right]\,. \end{aligned} \tag{4.49}$$

Now, from (4.47) and (4.49), we see that

$$\rho(z) = \hat{\hat{U}}(z, z_0)\,\hat{\hat{U}}_i(z, z_0)\,\rho(z_0)\hat{\hat{U}}_i^{\dagger}(z, z_0)\,\hat{\hat{U}}^{\dagger}(z, z_0)\,. \tag{4.50}$$

Hence, for the average of any observable O we have

$$\begin{aligned} \left\langle\hat{O}\right\rangle(z) &= \mathrm{Tr}\left\langle\rho(z)\hat{O}\right\rangle = \mathrm{Tr}\left\langle\hat{\hat{U}}(z, z_0)\,\hat{\hat{U}}_i(z, z_0)\,\rho(z_0)\hat{\hat{U}}_i^{\dagger}(z, z_0)\;.\hat{\hat{U}}^{\dagger}(z, z_0)\,\hat{O}\right\rangle \\ &= \mathrm{Tr}\left\langle\rho(z_0)\left\{\hat{\hat{U}}_i^{\dagger}(z, z_0)\,\hat{\hat{U}}^{\dagger}(z, z_0)\,\hat{O}\hat{\hat{U}}(z, z_0)\,\hat{\hat{U}}_i(z, z_0)\right\}\right\rangle\,. \end{aligned} \tag{4.51}$$

This equation (4.51) provides the general basic formula to compute the transfer map for $\left\langle\hat{O}\right\rangle$ across the system as will be seen below in the case of the present example.

Let us take z_0 and z to be respectively in the field-free input and output regions of the quadrupole magnet : $z_0 < z_{\rm n}\,,\; z > z_{\rm x}$. After some straightforward algebra we get

$$\hat{\hat{U}}(z, z_0) = \hat{\hat{U}}_F(z, z_{\rm x})\hat{\hat{U}}_L(z_{\rm x}, z_{\rm n})\hat{\hat{U}}_F(z_{\rm n}, z_0)\,,$$

$$\hat{\tilde{U}}_{i}(z, z_{o}) = \hat{\tilde{U}}_{i,F}(z, z_{x})\hat{\tilde{U}}_{i,L}(z_{x}, z_{n})\hat{\tilde{U}}_{i,F}(z_{n}, z_{o}) \equiv \hat{\tilde{U}}_{i,L}(z_{x}, z_{n}),$$

$$\hat{U}_{F}(z, z_{x}) = \exp\left\{\frac{i}{\hbar}\Delta z_{>}\left(p_{0} - \frac{1}{2p_{0}}\hat{p}_{\perp}^{2}\right)\right\}, \quad \text{with} \ \ \Delta z_{>} = z - z_{x},$$

$$\hat{U}_{L}(z_{x}, z_{n}) = \exp\left\{\frac{i}{\hbar}\ell\left[\left(p_{0} - \frac{1}{2p_{0}}\hat{p}_{\perp}^{2}\right) + \frac{1}{2}p_{0}K(x^{2} - y^{2})\right]\right\},$$

$$\text{with} \ \ K = qG/p_{0},$$

$$\hat{U}_{F}(z_{n}, z_{o}) = \exp\left\{\frac{i}{\hbar}\Delta z_{<}\left(p_{0} - \frac{1}{2p_{0}}\hat{p}_{\perp}^{2}\right)\right\}, \quad \text{with} \ \ \Delta z_{<} = z_{n} - z_{o},$$

$$\begin{aligned}
&\hat{\tilde{U}}_{i,L}(z_{x}, z_{n}) \\
&\quad = \exp\left\{-\frac{i}{\hbar}\eta\left[\left(\left(\frac{\sinh\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)p_{0}x + \left(\frac{\cosh\left(\sqrt{K}\,\ell\right) - 1}{K\ell}\right)\hat{p}_{x}\right)\sigma_{y}\right.\right. \\
&\qquad\qquad \left.\left. + \left(\left(\frac{\sin\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)p_{0}y - \left(\frac{\cos\left(\sqrt{K}\,\ell\right) - 1}{K\ell}\right)\hat{p}_{y}\right)\sigma_{x}\right]\right\}.
\end{aligned} \tag{4.52}$$

Now, using (4.51) and (4.52) the transfer maps for $\langle \boldsymbol{r}_{\perp}\rangle$ and $\langle \hat{\boldsymbol{p}}_{\perp}\rangle$ ($\equiv \langle \hat{\boldsymbol{\pi}}_{\perp}\rangle$ in this case) are obtained as follows : with $\lambda_{0} = h/p_{0}$, the de Broglie wavelength,

$$\begin{pmatrix} \langle x\rangle(z) \\ \langle \hat{p}_{x}\rangle(z)/p_{0} \\ \langle y\rangle(z) \\ \langle \hat{p}_{y}\rangle(z)/p_{0} \end{pmatrix} \approx \begin{pmatrix} T_{11}^{x} & T_{12}^{x} & 0 & 0 \\ T_{21}^{x} & T_{22}^{x} & 0 & 0 \\ 0 & 0 & T_{11}^{y} & T_{12}^{y} \\ 0 & 0 & T_{21}^{y} & T_{22}^{y} \end{pmatrix} \left(\begin{pmatrix} \langle x\rangle(z_{o}) \\ \langle \hat{p}_{x}\rangle(z_{o})/p_{0} \\ \langle y\rangle(z_{o}) \\ \langle \hat{p}_{y}\rangle(z_{o})/p_{0} \end{pmatrix}\right.$$

$$+\eta\begin{pmatrix} \left(\frac{\cosh\left(\sqrt{K}\,\ell\right)-1}{K\ell}\right)\langle\sigma_y\rangle(z_0) \\ -\left(\frac{\sinh\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)\langle\sigma_y\rangle(z_0) \\ -\left(\frac{\cos\left(\sqrt{K}\,\ell\right)-1}{K\ell}\right)\langle\sigma_x\rangle(z_0) \\ -\left(\frac{\sin\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)\langle\sigma_x\rangle(z_0) \end{pmatrix}\Bigg),$$

$$\begin{pmatrix} T^x_{11} & T^x_{12} \\ T^x_{21} & T^x_{22} \end{pmatrix} = \begin{pmatrix} 1 & \Delta z_> \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \cosh\left(\sqrt{K}\,\ell\right) & \frac{1}{\sqrt{K}}\sinh\left(\sqrt{K}\,\ell\right) \\ \sqrt{K}\sinh\left(\sqrt{K}\,\ell\right) & \cosh\left(\sqrt{K}\,\ell\right), \end{pmatrix}\begin{pmatrix} 1 & \Delta z_< \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} T^y_{11} & T^y_{12} \\ T^y_{21} & T^y_{22} \end{pmatrix} = \begin{pmatrix} 1 & \Delta z_> \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\left(\sqrt{K}\,\ell\right) & \frac{1}{\sqrt{K}}\sin\left(\sqrt{K}\,\ell\right) \\ -\sqrt{K}\,\sin\left(\sqrt{K}\,\ell\right) & \cos\left(\sqrt{K}\,\ell\right) \end{pmatrix}\begin{pmatrix} 1 & \Delta z_< \\ 0 & 1 \end{pmatrix},$$

$$\langle S_x\rangle(z) \approx \langle S_x\rangle(z_0) + \frac{4\pi\eta}{\lambda_0}\left(\left(\frac{\sinh\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)\langle xS_z\rangle(z_0) + \left(\frac{\cosh\left(\sqrt{K}\,\ell\right)-1}{K\ell p_0}\right)\langle\hat{p}_xS_z\rangle(z_0)\right),$$

$$\langle S_y\rangle(z) \approx \langle S_y\rangle(z_0) - \frac{4\pi\eta}{\lambda_0}\left(\left(\frac{\sin\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)\langle yS_z\rangle(z_0) - \left(\frac{\cos\left(\sqrt{K}\,\ell\right)-1}{K\ell p_0}\right)\langle\hat{p}_yS_z\rangle(z_0)\right),$$

$$\langle S_z\rangle(z) \approx \langle S_z\rangle(z_0) -$$

$$
\begin{aligned}
& \frac{4\pi\eta}{\lambda_0}\left\{\left(\frac{\sinh\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)\langle xS_z\rangle(z_o)\right. \\
& -\left(\frac{\sin\left(\sqrt{K}\,\ell\right)}{\sqrt{K}\,\ell}\right)\langle yS_y\rangle(z_o) \\
& +\left(\frac{\cosh\left(\sqrt{K}\,\ell\right)-1}{K\ell p_0}\right)\langle\hat{p}_xS_z\rangle(z_o) \\
& \left.+\left(\frac{\cos\left(\sqrt{K}\,\ell\right)-1}{K\ell p_0}\right)\langle\hat{p}_yS_y\rangle(z_o)\right\}.
\end{aligned}
\tag{4.53}
$$

So, we have got a fully quantum mechanical derivation of the traditional transfer map for the transverse phase-space, including the Stern-Gerlach effect (see [76]), in the case of a spin-$\frac{1}{2}$ particle beam propagating through a normal magnetic quadrupole lens : the lens is focusing (defocusing) in the yz-plane and defocusing (focusing) in the xz-plane when $K > 0$ ($K < 0$). The transverse Stern-Gerlach kicks to the trajectory slope ($\delta\langle\hat{\boldsymbol{p}}_\perp\rangle/p_0 \sim \eta$) are seen to disappear at relativistic energies, varying like $\sim 1/\gamma$. At nonrelativistic energies, with $\gamma \approx 1$, the kicks are $\sim G\ell\mu/m_0v^2$ where μ is the total magnetic moment. These results are in general agreement with the conclusions reached earlier [25, 76] based on semiclassical treatments. The spin map obtained above is seen to contain the paraxial Thomas-BMT map including the lowest order terms depending on $\boldsymbol{p}_\perp/p_0$. It should be also noted that the polarization transfer map is linear in the polarization components only when there is no spin-space correlation, i.e., for the classical behaviour to result one should have $\langle xS_z\rangle = \langle x\rangle\langle S_z\rangle$, $\langle yS_z\rangle = \langle y\rangle\langle S_z\rangle$, $\langle\hat{p}_xS_z\rangle = \langle\hat{p}_x\rangle\langle S_z\rangle$, etc..

4.3.2 Skew Magnetic Quadrupole

For a skew-magnetic-quadrupole lens the field is given by

$$
\boldsymbol{B} = (-G_s y, G_s x, 0) \tag{4.54}
$$

associated with the vector potential

$$\boldsymbol{A} = (0, 0, -G_s xy) \tag{4.55}$$

where G_s is assumed to be a constant in the lens region and zero outside

The basic accelerator optical Hamiltonian of the system is

$$\hat{H}(z) = \begin{cases} \hat{H}_F = -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2 , & \text{for } z < z_{\rm n} \text{ and } z > z_{\rm x} , \\ \hat{H}_L(z) = -p_0 + \frac{1}{2p_0}\hat{p}_\perp^2 + \frac{1}{2} q G_s xy - \frac{\eta_s p_0}{\ell}\left(x\sigma_y - y\sigma_x\right) , & \\ \qquad \text{for } z_{\rm n} \leq z \leq z_{\rm x} , \quad \text{with } \eta_s = (q + \gamma\epsilon) G_s \ell\hbar / 2p_0^2 . & \end{cases} \tag{4.56}$$

and

$$\begin{aligned} & \hat{\hat{U}}_{i,L}(z_{\rm x}, z_{\rm n}) \\ & = \exp\left\{ -\frac{\mathrm{i}}{\hbar}\frac{\eta_s}{2}\left[\left(-\frac{S^+(\ell)}{\sqrt{K_s}\,\ell} p_0 x + \frac{C^-(\ell)}{K_s \ell}\hat{p}_x + \frac{S^-(\ell)}{\sqrt{K_s}\,\ell} p_0 y + \frac{(C^+(\ell) - 2)}{K_s \ell}\hat{p}_y \right)\sigma_y \right.\right. \\ & \qquad \left.\left. - \left(-\frac{S^-(\ell)}{\sqrt{K_s}\,\ell} p_0 x + \frac{(C^+(\ell) - 2)}{K_s \ell}\hat{p}_x - \frac{S^+(\ell)}{\sqrt{K_s}\,\ell} p_0 y + \frac{C^-(\ell)}{K_s \ell}\hat{p}_y \right)\sigma_x \right] \right\} . \\ & \qquad C^{\pm}(z) = \cos\left(\sqrt{K_s}\, z\right) \pm \cosh\left(\sqrt{K_s}\, z\right) \\ & \qquad S^{\pm}(z) = \sin\left(\sqrt{K_s}\, z\right) \pm \sinh\left(\sqrt{K_s}\, z\right) \end{aligned} \tag{4.57}$$

Now, using (4.51) and (4.57) the transfer maps for $\langle \boldsymbol{r}_\perp \rangle$ and $\langle \hat{\boldsymbol{p}}_\perp \rangle$ ($\equiv \langle \hat{\boldsymbol{\pi}}_\perp \rangle$) are obtained as follows : with $\lambda_0 = h/p_0$, the de Broglie wavelength,

$$\begin{pmatrix} \langle x \rangle (z) \\ \langle \hat{p}_x \rangle (z)/p_0 \\ \langle y \rangle (z) \\ \langle \hat{p}_y \rangle (z)/p_0 \end{pmatrix} \approx T \left(\begin{pmatrix} \langle x \rangle (z_0) \\ \langle \hat{p}_x \rangle (z)/p_0 \\ \langle y \rangle (z_0) \\ \langle \hat{p}_y \rangle (z)/p_0 \end{pmatrix} \right.$$

$$+\frac{\eta_s}{2}\left(\begin{array}{c}\frac{-C^-(\ell)\langle\sigma_y\rangle(z_0)+\left(C^+(\ell)-2\right)\langle\sigma_z\rangle(z_0)}{K_s\ell}\\ \frac{-S^+(\ell)\langle\sigma_y\rangle(z_0)+S^-(\ell)\langle\sigma_z\rangle(z_0)}{\sqrt{K_s}\ell}\\ \frac{-\left(C^+(\ell)-2\right)\langle\sigma_y\rangle(z_0)+C^-(\ell)\langle\sigma_z\rangle(z_0)}{K_s\ell}\\ \frac{-S^-(\ell)\langle\sigma_y\rangle(z_0)+S^+(\ell)\langle\sigma_z\rangle(z_0)}{\sqrt{K_s}\ell}\end{array}\right)\Bigg),$$

$$T = M_> M M_<\,,$$

$$M = \frac{1}{2}\begin{pmatrix} C^+(\ell) & \frac{S^+(\ell)}{\sqrt{K_s}} & C^-(\ell) & \frac{S^-(\ell)}{\sqrt{K_s}} \\ -\sqrt{K_s}S^-(\ell) & C^+(\ell) & -\sqrt{K_s}S^+(\ell) & C^-(\ell) \\ C^-(\ell) & \frac{S^-(\ell)}{\sqrt{K_s}} & C^+(\ell) & \frac{S^+(\ell)}{\sqrt{K_s}} \\ -\sqrt{K_s}S^+(\ell) & C^-(\ell) & -\sqrt{K_s}S^-(\ell) & C^+(\ell) \end{pmatrix}$$

$$M_{\lessgtr} = \begin{pmatrix} 1 & \Delta z_{\lessgtr} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta z_{\lessgtr} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.58}$$

In the thin lens approximation the transfer matrix M simplifies to

$$\begin{aligned} M &= \begin{pmatrix} 1 & \ell & 0 & 0 \\ 0 & 1 & -1/f & 0 \\ 0 & 0 & 1 & \ell \\ -1/f & 0 & 0 & 1 \end{pmatrix} \\ 1/f &= K_s\ell\,. \end{aligned} \tag{4.59}$$

Along the diagonal the thin lens matrix blocks look like drift through space of length ℓ while the off-diagonal matrix blocks describe the focusing of the skew quadrupole lens, in agreement with the well-known classical results [49].

For spin, the transfer map reads

$$\langle S_x\rangle(z) \approx \langle S_x\rangle(z_0) +$$

$$
\frac{2\pi\eta_s}{\lambda_0}\left\{-\left(\frac{S^+(\ell)}{\sqrt{K_s}\,\ell}\right)\langle xS_z\rangle(z_0)+\left(\frac{C^-(\ell)}{K_s\ell p_0}\right)\langle\hat{p}_xS_z\rangle(z_0)\right.
$$
$$
\left.-\left(\frac{S^-(\ell)}{\sqrt{K_s}\,\ell}\right)\langle yS_z\rangle(z_0)+\left(\frac{C^+(\ell)-2}{K_s\ell p_0}\right)\langle\hat{p}_yS_z\rangle(z_0)\right\},
$$

$$
\begin{aligned}
\langle S_y\rangle(z)\approx\ &\langle S_y\rangle(z_0)+\\
&\frac{2\pi\eta_s}{\lambda_0}\left\{-\left(\frac{S^-(\ell)}{\sqrt{K_s}\,\ell}\right)\langle xS_z\rangle(z_0)+\left(\frac{C^+(\ell)-2}{K_s\ell p_0}\right)\langle\hat{p}_xS_z\rangle(z_0)\right.\\
&\left.-\left(\frac{S^+(\ell)}{\sqrt{K_s}\,\ell}\right)\langle yS_z\rangle(z_0)+\left(\frac{C^-(\ell)}{K_s\ell p_0}\right)\langle\hat{p}_yS_z\rangle(z_0)\right\},
\end{aligned}
$$

$$
\begin{aligned}
\langle S_z\rangle(z)\approx\ &\langle S_z\rangle(z_0)\\
&+\frac{2\pi\eta_s}{\lambda_0}\left\{\left\{\left(\frac{S^+(\ell)}{\sqrt{K_s}\,\ell}\right)\langle xS_x\rangle(z_0)-\left(\frac{C^-(\ell)}{K_s\ell p_0}\right)\langle\hat{p}_xS_x\rangle(z_0)\right.\right.\\
&\left.+\left(\frac{S^-(\ell)}{\sqrt{K_s}\,\ell}\right)\langle yS_x\rangle(z_0)-\left(\frac{C^+(\ell)-2}{K_s\ell p_0}\right)\langle\hat{p}_yS_x\rangle(z_0)\right\}\\
&+\left\{-\left(\frac{S^-(\ell)}{\sqrt{K_s}\,\ell}\right)\langle xS_y\rangle(z_0)+\left(\frac{C^+(\ell)-2}{K_s\ell p_0}\right)\langle\hat{p}_xS_y\rangle(z_0)\right.\\
&\left.\left.-\left(\frac{S^+(\ell)}{\sqrt{K_s}\,\ell}\right)\langle yS_y\rangle(z_0)+\left(\frac{C^-(\ell)}{K_s\ell p_0}\right)\langle\hat{p}_yS_y\rangle(z_0)\right\}\right\},
\end{aligned}
\tag{4.60}
$$

4.4 Stern-Gerlach Force

Using the general theory, let us now understand the longitudinal Stern-Gerlach kicks [25] in a general inhomogeneous magnetic field. For $\hat{\pi}_z = -(\hat{H}^{(A)} + qA_z)$ we get, from (4.30),

$$
\begin{aligned}
\frac{d}{dz}\langle\hat{\pi}_z\rangle &= \left\{\frac{\mathrm{i}}{\hbar}\left\langle\left[\hat{H}^{(A)}+qA_z\,,\,\hat{H}^{(A)}\right]\right\rangle-\left\langle\frac{\partial}{\partial z}\left(\hat{H}^{(A)}+qA_z\right)\right\rangle\right\}\\
&= \left\langle\frac{\mathrm{i}}{\hbar}\left[qA_z\,,\,\hat{H}^{(A)}\right]-\frac{1}{2p_0}\frac{\partial}{\partial z}\hat{\pi}_\perp^2\right\rangle-\frac{\gamma m_0}{p_0}\left\langle\frac{\partial}{\partial z}\left(\underline{\boldsymbol{\Omega}}_s\cdot\boldsymbol{S}\right)\right\rangle\\
&= \frac{q}{p_0}\left\langle\frac{1}{2}\left(\hat{\boldsymbol{\pi}}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\boldsymbol{\pi}}\right)_z\right\rangle-\frac{\gamma m_0}{p_0}\left\langle\frac{\partial}{\partial z}\left(\underline{\boldsymbol{\Omega}}_s\cdot\boldsymbol{S}\right)\right\rangle
\end{aligned}
$$

$$= \frac{q}{p_0}\left\langle\frac{1}{2}\left(\hat{\pi}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\pi}\right)_z\right\rangle + \frac{1}{p_0}\left\langle\frac{\partial}{\partial z}\left\{(q+\epsilon)B_zS_z+(q+\gamma\epsilon)\boldsymbol{B}_\perp\cdot\boldsymbol{S}_\perp\right\}\right\rangle . \tag{4.61}$$

The first term on the r.h.s. of (4.61) corresponds to the Lorentz force and the rest of it corresponds to the Stern-Gerlach force due to the longitudinal gradient of the field (*i.e.*, gradient in the z-direction). This is easily recognized by multiplying both sides of (4.61) by $v_z \approx p_0/\gamma m_0$ and comparing the resulting equation for $\frac{d}{dt}\langle\hat{\pi}_z\rangle$ with the classical equation of motion for π_z as is done in the case of $\hat{\boldsymbol{\pi}}_\perp$ in (4.36). Collecting together (4.35) and (4.61) we get for the z-evolution of $\langle\hat{\boldsymbol{\pi}}\rangle$

$$\frac{d}{dz}\langle\hat{\boldsymbol{\pi}}\rangle \approx \frac{q}{p_0}\left\langle\frac{1}{2}\left(\hat{\boldsymbol{\pi}}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\boldsymbol{\pi}}\right)\right\rangle + \frac{1}{p_0}\left\langle\boldsymbol{\nabla}\left\{(q+\epsilon)B_zS_z+(q+\gamma\epsilon)\boldsymbol{B}_\perp\cdot\boldsymbol{S}_\perp\right\}\right\rangle . \tag{4.62}$$

For any given field configuration $\boldsymbol{B}$, with a specified $\boldsymbol{A}$, the solution of this equation (4.62) is given by

$$\langle\hat{\boldsymbol{\pi}}\rangle(z) = \mathrm{Tr}\left\langle\rho(z_0)\hat{U}^\dagger(z,z_0)\hat{\boldsymbol{\pi}}\hat{U}(z,z_0)\right\rangle , \quad \text{for any } z > z_0 , \tag{4.63}$$

and hence the spin-dependent Stern-Gerlach kick to the kinetic momentum and the resultant spin-dependent splitting of the kinetic energy at any $z > z_0$ can be calculated.

Multiplying both sides of (4.62) by $v_z \approx p_0/\gamma m_0$, it follows that

$$\begin{aligned}\frac{d}{dt}\langle\hat{\boldsymbol{\pi}}\rangle &\approx \frac{q}{\gamma m_0}\left\langle\frac{1}{2}\left(\hat{\boldsymbol{\pi}}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\boldsymbol{\pi}}\right)\right\rangle + \frac{1}{\gamma m_0}\left\langle\boldsymbol{\nabla}\left\{q\boldsymbol{B}\cdot\boldsymbol{S}+\epsilon\left(B_zS_z+\gamma\boldsymbol{B}_\perp\cdot\boldsymbol{S}_\perp\right)\right\}\right\rangle \\ &= \frac{q}{\gamma m_0}\left\langle\frac{1}{2}\left(\hat{\boldsymbol{\pi}}\times\boldsymbol{B}-\boldsymbol{B}\times\hat{\boldsymbol{\pi}}\right)\right\rangle - \left\langle\boldsymbol{\nabla}\left(\boldsymbol{\Omega}_s\cdot\boldsymbol{S}\right)\right\rangle ,\end{aligned} \tag{4.64}$$

in which the first term represents the Lorentz force and the second term represents the Stern-Gerlach force. This equation (4.64) for orbital motion of a Dirac particle moving predominantly along the z-direction is seen to account, under the paraxial approximation, for both the Lorentz and the Stern-Gerlach forces. It may be noted that our formalism facilitates the computation of the transfer maps for the beam observables over any interval (z_0, z) along the axis by the use of direct z-evolution formulae, like in (4.63), and we are considering the time evolution equations such as (4.64) only for the sake of comparison with the classical equations of motion. In the instantaneous rest frame of the particle with $\gamma = 1$ the second term in (4.64) is seen to correspond to the familiar Stern-Gerlach force

$$\boldsymbol{F}_{SG} = -\boldsymbol{\nabla} U\,, \quad U = -\mu\boldsymbol{\sigma} \cdot \boldsymbol{B}\,, \tag{4.65}$$

where μ is the total magnetic moment of the particle; note that in (4.65), apart from the spin, the field components, the coordinates, etc., are also defined in the rest frame of the particle.

It is of interest to know the relative merits and demerits of spin-splitter devices employing the transverse and longitudinal Stern-Gerlach kicks. When the fields $\boldsymbol{B}$ in such devices are known explicitly one can directly use the formula (4.63) for such a study. But, to have an idea of the situation in a general context, one can use the standard classical relativistic dynamics [81, 82] starting with the form of the Stern-Gerlach force (4.65) which has been understood on the basis of the Dirac equation; the result has to agree with the classical limit of the quantum mechanical computation. Such a study [25] based on classical relativistic dynamics seems to suggest that, at high energies, devices employing the longitudinal kick are more favourable than those employing the transverse kick. To be more precise, with G_z denoting the longitudinal magnetic gradient $\frac{\partial B_z}{\partial z}$ active over a region of length L in a device employing the longitudinal kick, the fractional increase in the longitudinal momentum, $\delta p_z/p_0$, turns out to be $G_z\mu L/m_0 v^2$, which becomes almost independent of energy as γ increases (see [25] for details of the calculation). In the case of a device employing the transverse kick, the fractional increase in the transverse momentum varies like $\sim 1/\gamma$ and thus decreases as γ increases, as we have seen above in the example of the quadrupolar magnetic field (see [25] for details of the calculation based on classical relativistic dynamics). Thus, one can conclude generally that at high energies a spin splitter with longitudinal kick should be more favourable than one with transverse kick, leaving aside all technical details such as the practical realization of the required longitudinal magnetic gradient and the way of exploiting the attained spin-dependent energy spread. At lower energies, the kicks are larger in both the cases.

In summary, we have demonstrated how one can obtain a fully quantum mechanical understanding of the accelerator beam optics for a spin-$\frac{1}{2}$ particle, with anomalous magnetic moment, starting *ab initio* from the Dirac-Pauli equation. To this end, we have used a beam optical representation of the Dirac theory, following [8]-[12], and have shown that such an approach, in the lowest order approximation, leads naturally to a picture of orbital and spin dynamics based on the Lorentz force, the Stern-Gerlach force and the Thomas-BMT equation for spin evolution, as is to be expected. Only the lowest order (paraxial) approximation has been considered in detail. To illustrate the general theory we have considered the computation of the transfer maps for the spin components and the transverse phase-space, including the transverse Stern-Gerlach kicks, in the case of a normal magnetic quadrupole lens, and a brief understanding of the longitudinal Stern-Gerlach kicks in a general inhomogeneous magnetic field. It is found that the above theory supports the spin-splitter concepts based on transverse and longitudinal Stern-Gerlach kicks [74]-[80],[25]. It is clear from the general theory, presented briefly here, that the approach is suitable to handle any magnetic optical element with straight axis and computations can be carried out to any order of accuracy desired by easily extending the order of approximation. In fact, even the lowest order approximation reveals the nature of deviations from the classical behaviour for spin evolution, namely, the dependence on differences $\langle xS_z\rangle - \langle x\rangle \langle S_z\rangle$, $\langle yS_z\rangle - \langle y\rangle \langle S_z\rangle$, $\langle p_xS_z\rangle - \langle p_x\rangle \langle S_z\rangle$, etc..

The suggestion for using the Stern-Gerlach kicks to produce high energy polarized beams ([25] and references therein) has aroused much interest in the exact form of the Stern-Gerlach force in the relativistic region; there has been ambiguities about it in the literature. This question is being thoroughly analysed recently ([71] and references therein). From the exhaustive analysis presented in [71] it seems that this question, which also involves the problem of proper choice of the position operator in the context of relativistic quantum theory, can be settled only through suitable experimentation.

Chapter 5

Phase-space formalism of the quantum theory of charged-particle beam optics

In this chapter we indicate an alternate approach to the quantum theory of charged-particle beam transport based on the Wigner phase-space distributions. Such an approach would provide a link between classical and quantum descriptions [26] (see also [5, 27] and [83]-[86] for works related to the use of Wigner distribution in charged-particle optics). The present chapter is confined to the scalar case in paraxial approximation. The possibility of extending the phase-space formalism to the study of aberrating systems and the Dirac, or spinor, charged-particle beam optics is also briefly noted.

We start with the beam-optical form of the Schrödinger/Klein-Gordon equation (2.64) for the z-evolution of the paraxial beam wavefunction $\Psi(\boldsymbol{r}_\perp, z)$:

$$\begin{aligned} i\hbar\frac{\partial\Psi}{\partial z} &= \hat{\mathbf{H}}_o\Psi\,, \\ \hat{\mathbf{H}}_o &= -p_0 - qA_z + \frac{1}{2p_0}\left(\hat{\pi}_\perp^2 + \hat{p}^2\right) \end{aligned} \tag{5.1}$$

We shall cite the example of the axially symmetric magnetic lens or the magnetic round lens to illustrate the use of Wigner functions in charged-particle optics. The round lens in paraxial approximation is described by the vector potential

$$\boldsymbol{A} = \left(-\frac{1}{2}yB(z)\,,\,\frac{1}{2}xB(z)\,,\,0\right)\,. \tag{5.2}$$

After dropping the constant p_0, equation (5.1) becomes

$$\mathrm{i}\hbar\frac{\partial\Psi(\boldsymbol{r}_\perp,z)}{\partial z} = \left[\frac{1}{2p_0}\left(\hat{p}_\perp^2 + p_0^2\alpha^2\hat{r}_\perp^2\right) - \alpha(z)\hat{L}_z\right]\Psi(\boldsymbol{r}_\perp,z)\,\cdot \tag{5.3}$$

where $\alpha = \frac{qB(z)}{2p_0}$ and $\hat{L}_z$ is the z-component of the angular momentum. Since equation (5.3) is invariant under rotation around the z-axis, we can make a transformation of the wavefunction Ψ to another, say $\psi(\boldsymbol{r}_\perp,z)$, referred to the rotating coordinate system along the z-axis. To this end, we let

$$\psi(\boldsymbol{r}_\perp,z) = \exp\left(\frac{\mathrm{i}}{\hbar}\theta(z)\hat{L}_z\right)\Psi(\boldsymbol{r}_\perp,z), \tag{5.4}$$

where $\theta(z) = \int_{z_{\rm in}}^{z}\alpha(z)$. Then, for the paraxial beam wavefunction $\psi(\boldsymbol{r}_\perp,z)$, referred to the rotating coordinate system along the z-axis, we obtain the Glaser equation [4].

$$\mathrm{i}\hbar\frac{\partial\psi(\boldsymbol{r}_\perp,z)}{\partial z} = \left[\frac{1}{2p_0}\left(\hat{p}_\perp^2 + p_0^2\alpha^2\hat{r}_\perp^2\right)\right]\psi(\boldsymbol{r}_\perp,z)\,. \tag{5.5}$$

As we shall see soon, this equation explains the focusing of the beam along both the x and y directions; the disappearance of the $\hat{L}_z$-term from (5.3) under the transformation in (5.4) explains the quantum mechanics of the image rotation as is seen from the inverse map to the original untwisted coordinate system, $\Psi = \exp\left(-\frac{\mathrm{i}}{\hbar}\theta(z)\hat{L}_z\right)\psi$.

Recollect, that equation (5.5) is similar to Schrödinger's equation for a two-dimensional harmonic oscillator with time-dependent frequency. For a specific time-dependence one can solve the harmonic oscillator problem analytically or otherwise. For our purpose we do not need the eigenfunctions and the eigenvalues. We are interested in obtaining closed form expressions for the optical parameters, like the focal length, in terms of the design parameters, p_0, $B(z)$, etc. To this end, we require the z-propagator for (5.5) to obtain the relation

$$|\psi(z_{\text{out}})\rangle = \hat{G}(z_{\text{out}}, z_{\text{in}})\,|\psi(z_{\text{in}})\rangle\,, \tag{5.6}$$

connecting the states of the beam in the input ($z \leq z_{\text{in}}$) and the output ($z \geq z_{\text{out}}$) field-free regions of the lens system. More explicitly, we want the relation

$$\psi(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}) = \int d^2\boldsymbol{r}_{\perp,\text{in}}\, G(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}; \boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}})\,\psi(\boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}})\,, \tag{5.7}$$

with $G(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}; \boldsymbol{r}_{\perp,\text{out}}, z_{\text{in}}) = \left\langle \boldsymbol{r}_{\perp,\text{out}} \left| \hat{G}(z_{\text{out}}, z_{\text{in}}) \right| \boldsymbol{r}_{\perp,\text{in}} \right\rangle$. To obtain the required Green's function G we can utilize the well-known result (see, e.g., [40]) that the time-evolution operator, or the Feynman propagator, for the most general time-dependent quadratic Hamiltonian, can be be elegantly expressed in a closed form in terms of two linearly independent solutions to the classical equations of motion (see also [4] for the pioneering related work of Glaser in the context of electron optics). Now, our equations of motion are the trajectory equations

$$\boldsymbol{r}_\perp''(z) + \alpha(z)^2 \boldsymbol{r}_\perp(z) = 0\,, \tag{5.8}$$

Let the two linearly independent solutions of (5.8), identical for $\boldsymbol{r}_\perp = x$ or y, be taken, as is usual, to be $h(z)$ and $g(z)$ with the initial conditions

$$g(z_{\text{in}}) = h'(z_{\text{in}}) = 1\,, \qquad g'(z_{\text{in}}) = h(z_{\text{in}}) = 0\,. \tag{5.9}$$

Then the required Green's function is given by

$$G(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}; \boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}) = G(x_{\text{out}}, z_{\text{out}}; x_{\text{in}}, z_{\text{in}})\, G(y_{\text{out}}, z_{\text{out}}; y_{\text{in}}, z_{\text{in}})\,, \tag{5.10}$$

with, for $\boldsymbol{r}_\perp = x$ and y,

$$G(\boldsymbol{r}_\perp, z; \boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}) = \sqrt{\frac{p_0}{i\hbar 2\pi h(z)}}\exp\left\{\frac{ip_0}{2\hbar h(z)}\left[g(z)\boldsymbol{r}_{\perp,\text{in}}^2 + h'(z)\boldsymbol{r}_\perp^2 - 2\boldsymbol{r}_\perp \boldsymbol{r}_{\perp,\text{in}}\right]\right\}. \tag{5.11}$$

At this stage we introduce the Wigner function, to obtain the transfer maps.

The Wigner phase-space distribution function [87] associated with a quantum mechanical wavefunction $\psi(q)$ is given by

$$\mathcal{W}_\psi(q,p_q) := \frac{1}{2\pi\hbar}\int d\sigma\,\psi^*\left(q-\frac{\sigma}{2}\right)\psi\left(q+\frac{\sigma}{2}\right)\exp\left(-\mathrm{i}\frac{\sigma p_q}{\hbar}\right) \qquad (5.12)$$

where we have considered the one dimensional case and denoted by p_q the momentum canonically conjugate to q. Wigner functions are always real but not necessarily positive. The Wigner, and other phase-space distributions later considered, have been extensively studied and applied (see, e.g., [88] and [89]–[94]). Here, we shall recall a few relevant properties. If $\hat{\mathcal{O}}\,(q,\hat{p}_q)$ is the quantum mechanical operator associated with the classical observable $\mathcal{O}\,(q,p_q)$, in accordance with the Weyl correspondence rule, then

$$\left\langle\psi\left|\hat{\mathcal{O}}\right|\psi\right\rangle = \int\int dq\,dp_q\,\mathcal{W}_\psi\,(q,p_q)\,\mathcal{O}\,(q,p_q)\,, \quad \int\int dq\,dp_q\,\mathcal{W}\,(q,p_q) = 1\,, \quad (5.13)$$

as if the Wigner function is a classical phase-space probability distribution function which it is not. Further, as should be expected,

$$\int dp_q\,\mathcal{W}_\psi\,(q,p_q) = |\psi(q)|^2\,, \quad \int dq\,\mathcal{W}_\psi\,(q,p_q) = |\,\langle p_q|\psi\rangle\,|^2\,, \qquad (5.14)$$

where $\langle p_q|\psi\rangle$ is the wavefunction in the momentum(p_q) representation.

It is clear that one can express the Wigner function $\mathcal{W}_{\rm out}(q_{\rm out},p_{q,\rm out})$ in the *output* plane in terms of the Wigner function $\mathcal{W}_{\rm in}(q_{\rm in},p_{q,\rm in})$ in the *input* plane by combining (5.7) and (5.12) (see, e.g., [95]). Thus one can write

$$\mathcal{W}_{\rm out}(q_{\rm out},p_{q,\rm out}) = \int\int dq_{\rm in}\,dp_{q,\rm in}\,K\,(q_{\rm out},p_{q,\rm out};q_{\rm in},p_{q,\rm in})\,\mathcal{W}_{\rm in}(q_{\rm in},p_{q,\rm in})\,, \quad (5.15)$$

where the function $K\,(q_{\rm out},p_{q,\rm out};q_{\rm in},p_{q,\rm in})$ is completely determined by the system, according to the relation giving the Wigner transform of $G(q,z,q_{\rm in},z_{\rm in};p)$:

$$K\,(q,p_q;q_{\rm in},p_{q,\rm in}) \;:=\; \mathcal{W}_G(q,p_q;q_{\rm in},p_{q,\rm in})$$

$$
\begin{aligned}
&= \frac{1}{2\pi\hbar}\int\int d\sigma\, d\sigma_{\rm in} G^*\left(q-\frac{\sigma}{2}; q_{\rm in}-\frac{\sigma_{\rm in}}{2}\right) G\left(q+\frac{\sigma}{2}; q_{\rm in}+\frac{\sigma_{\rm in}}{2}\right) \\
&\qquad \times \exp\left[-\frac{i}{\hbar}\left(p_q\sigma - p_{q,{\rm in}}\sigma_{\rm in}\right)\right] \\
&= \delta\left[q_{\rm in} - h'(z)q + h(z)p_q/p\right]\delta\left[p_{q,{\rm in}} + pg'(z)q - g(z)p_q\right] \\
&= \delta\left[q_{\rm in} - Dq + Bp_q\right]\delta\left[p_{q,{\rm in}} + Cq - Ap_q\right] . \qquad (5.16)
\end{aligned}
$$

The function $K\left(q, p_q; q_{\rm in}, p_{q,{\rm in}}\right)$ is the ray spread function corresponding to the response of the system in the space-momentum domain to the input signal $\mathcal{W}_{\rm in}(q_{\rm in}, p_{q,{\rm in}})$ and defines the input-output relationship in the phase-space. The four real constants, A, B, C and D, such that $AD - BC = 1$, constitute the ray-transfer matrix, $\mathcal{T}$. This result obtained for the magnetic lens is a general feature of any first-order system: any first-order system can be specified through such a ray-transfer matrix, or an ABCD-matrix,

$$
\mathcal{T} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad AD - BC = 1 , \qquad \mathcal{T} \in SL\left(2, \Re\right) , \qquad (5.17)
$$

an element of the symplectic group of transformations;

$$
\tilde{\mathcal{T}}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\mathcal{T} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \qquad (5.18)
$$

(In general, the transformation of the 4-dimensional phase-space $(x, p_x; y, p_y)$ will be described by an $SL\left(4, \Re\right)$ matrix for a linear system). We have considered the case of a single degree of freedom. Hence, it is clear from (5.15) and (5.16) that, for first-order charged-particle optical systems, the Wigner distribution function has the elegant property that

$$
\mathcal{W}\left((q, p_q); z\right) = \mathcal{W}_{\rm in}\left((q, p_q)\tilde{\mathcal{T}}^{-1}; z_{\rm in}\right) , \qquad (5.19)
$$

when the motion in the x and y planes are decoupled. (This is completely analogous to the similar result in the (photon) optical case (see, e.g., [96] and [95]).

The explicit *input-output* relation in the phase-space is seen to be

$$
\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm out} = \mathcal{T}\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} = \begin{pmatrix} g(z) & h(z) \\ g'(z) & h'(z) \end{pmatrix}_{\rm out}\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} . \qquad (5.20)
$$

Only at this stage we need the specific solutions, $g(z)$ and $h(z)$, for the classical trajectory equations (5.8), in order to compute (A, B, C, D). For the problem at hand one may obtain the two solutions of (5.8) by a method that does not destroy the required symplecticity inherent in the transfer matrix $\mathcal{T}$. As before we can use the series method described in detail in Chapter II (see 2.161–2.164) or equivalently we may use the method of Picard and Lindelöf [97]. For a thin lens, to a required degree of approximation, one gets

$$\begin{aligned}\begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm out} &= \begin{pmatrix} 1 - \frac{z_{\rm out}}{f} & \left(z_{\rm out} - z_{\rm in} + \frac{z_{\rm out} z_{\rm in}}{f}\right) \\ -\frac{1}{f} & 1 + \frac{z_{\rm in}}{f} \end{pmatrix} \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} \\ &= \begin{pmatrix} 1 & z_{\rm out} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & -z_{\rm in} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{r}_\perp \\ \boldsymbol{p}_\perp/p_0 \end{pmatrix}_{\rm in} .\end{aligned} \quad (5.21)$$

where

$$\frac{1}{f} \approx \int_{z_{\rm in}}^{z_{\rm out}} dz\, \alpha(z)^2 \quad (5.22)$$

Then, it is straightforward to see that the system behaves as a convergent lens with the focal length f; the thin lens approximation implies $z_{\rm out} - z_{\rm in} \ll f$. The focal length thus derived under the paraxial approximation coincides with the classical formula due to Busch ([43, 44]). So, we have a quantum mechanical derivation of the focal length of a thin magnetic lens for a paraxial charged-particle beam. It may be noted that in the field-free regions outside the lens system there is no difference between the kinetic momentum and the canonical momentum.

An alternative way of understanding the result in (5.19) is to use an operator approach (see references [92]–[94] and [98]–[100], for the related aspects) by which one has

$$\mathcal{W}_\rho(q, p_q) = \frac{1}{\pi} {\rm Tr}\left[\hat{\rho} \hat{D}_{q,p_q} \hat{\mathsf{P}} \hat{D}^\dagger_{q,p_q}\right] , \quad (5.23)$$

where $\hat{\rho}$ is the density operator of the system, $\hat{\mathsf{P}}$ is the parity operator (see, [93] and [98]– [100] for more details) and

$$\hat{D}_{q,p_q} = \exp\left[-\frac{\rm i}{\hbar}(q, p_q)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \hat{q} \\ \hat{p}_q \end{pmatrix}\right] . \quad (5.24)$$

We have considered above the pure states ($\hat{\rho} = |\psi\rangle\langle\psi|$). From (5.6) it is clear that

$$\hat{\rho}_{\rm out} = \hat{G}\left(z_{\rm out}, z_{\rm in}\right)\hat{\rho}_{\rm in}\hat{G}^{\dagger}\left(z_{\rm out}, z_{\rm in}\right). \tag{5.25}$$

Then, in general,

$$\begin{aligned} \mathcal{W}_{\rho_{\rm out}} &= \frac{1}{\pi}{\rm Tr}\left[\hat{G}\rho_{\rm in}\hat{G}^{\dagger}\hat{D}_{q,p_q}\hat{\rm P}\hat{D}^{\dagger}_{q,p_q}\right] \\ &= \frac{1}{\pi}{\rm Tr}\left[\rho_{\rm in}\hat{G}^{\dagger}\hat{D}_{q,p_q}\hat{G}\hat{G}^{\dagger}\hat{\rm P}\hat{G}\hat{G}^{\dagger}\hat{D}^{\dagger}_{q,p_q}\hat{G}\right], \end{aligned} \tag{5.26}$$

since $\hat{G}$ is unitary in the present case; integrating (5.5), formally, we have

$$\begin{aligned} \hat{G}\left(z_{\rm out}, z_{\rm in}\right) &= \wp\left\{\exp\left[-\frac{\rm i}{\hbar}\int_{z_{\rm in}}^{z_{\rm out}} dz\,\left(\frac{1}{2p_0}\left(\hat{p}_\perp^2 + \alpha^2 r_\perp^2\right)\right)\right]\right\} \\ &\approx \exp\left[-\frac{\rm i}{\hbar}\left(\frac{(z_{\rm out}-z_{\rm in})}{2p_0}\hat{p}_q^2 + \frac{p_0}{2f}\hat{q}^2\right)\right], \end{aligned} \tag{5.27}$$

where $\wp$ stands for the z-ordering of the exponential. Note that this $\hat{G}$ commutes with the parity operation. Thus we get

$$\mathcal{W}_{\rho_{\rm out}} = \frac{1}{\pi}{\rm Tr}\left[\rho_{\rm in}\hat{G}^{\dagger}\hat{D}_{q,p_q}\hat{G}\hat{\rm P}\hat{G}^{\dagger}\hat{D}^{\dagger}_{q,p_q}\hat{G}\right]. \tag{5.28}$$

Now let us note that, under the approximation considered,

$$\hat{G}^{\dagger}\begin{pmatrix}\hat{q}\\ \hat{p}_q\end{pmatrix}\hat{G} = \begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}\hat{q}\\ \hat{p}_q\end{pmatrix}. \tag{5.29}$$

In view of the symplecticity of $\mathcal{T}$, we get

$$\begin{aligned} \hat{G}^{\dagger}\hat{D}_{q,p_q}\hat{G} &\approx \exp\left[-\frac{\rm i}{\hbar}\left((q,p_q)\mathcal{T}^{-1}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\hat{q}\\ \hat{p}_q\end{pmatrix}\right)\right], \\ \hat{G}^{\dagger}\hat{D}^{\dagger}_{q,p_q}\hat{G} &= \left(\hat{G}^{\dagger}\hat{D}_{q,p_q}\hat{G}\right)^{\dagger}. \end{aligned} \tag{5.30}$$

Thus, equation (5.26) leads to the result in (5.19) in view of (5.30) (see, e.g., [96] for analogy with the (photon) optical case). Similar results for the Wigner function of electron beam transport through a linear element has been obtained by Dattoli *et al.* [101] using their formalism based on a Schrödinger-like equation. Actually, in (5.26) we have the transformation law of the Wigner function in the general situation when the paraxial and thin lens approximations may not be strictly valid. In such a case $\hat{G}$ will contain terms of higher power of $(\hat{q}, \hat{p}_q)$ corresponding to the various aberrations. Very recently we learned about the work by Fedele *et al.*,[27] where they have used the Wigner functions beyond the paraxial approximation.

In the classical (geometrical electron optics) limit ($\hbar \to 1$) the map in (5.29) becomes the classical Lie map in the phase-space of the system (see [15]- [21] for the pioneering development of Lie algebraic tools for classical charged-particle optics by Dragt *et al.*). Since the Wigner function is the quantum mechanical analogue of the classical phase-space density one can hope that the study of the evolution of the beam Wigner function along the beam optical axis of the system would lead to a better understanding of the well-known classical description of charged-particle optics and would provide quantum corrections whenever necessary.

To conclude, we would like to make some remarks. First, we shall see how the Green's function $G\left(z_{\rm out}, z_{\rm in}\right)$ in (5.7) can be obtained using an alternative approach. As already seen above (see (5.27)), up to the lowest order approximation, we have for the thin lens

$$\hat{G} \approx \exp\left[-{\rm i}\frac{\Delta z}{\hbar}\hat{\rm H}\right], \qquad \hat{\rm H} = \frac{\hat{p}_\perp^2}{2p_0} + \frac{p_0}{2(\Delta z)f}{}_{\perp}^{\;2}, \qquad \Delta z = \left(z_{\rm out} - z_{\rm in}\right), \qquad (5.31)$$

where $\hat{\rm H}$ is similar to the Hamiltonian of a two dimensional harmonic oscillator of mass p_0 and circular frequency $\omega = \sqrt{\frac{1}{(\Delta z)f}}$. The Green's function required is just $\left\langle \boldsymbol{r}_{\perp,{\rm out}} \middle| \hat{G}\left(z_{\rm out}, z_{\rm in}\right) \middle| \boldsymbol{r}_{\perp,{\rm in}} \right\rangle$ which may be expressed as

$$\begin{aligned}
&\left\langle \boldsymbol{r}_{\perp,{\rm out}} \middle| \hat{G}\left(z_{\rm out}, z_{\rm in}\right) \middle| \boldsymbol{r}_{\perp,{\rm in}} \right\rangle \\
&= \sum_{(n_1,n_2)} \sum_{(m_1,m_2)} \left\langle \boldsymbol{r}_{\perp,{\rm out}} | n_1, n_2 \right\rangle \left\langle n_1, n_2 \middle| \hat{G}\left(z_{\rm out}, z_{\rm in}\right) \middle| m_1, m_2 \right\rangle \left\langle m_1, m_2 | \boldsymbol{r}_{\perp,{\rm in}} \right\rangle \\
&= \sum_{(n_1,n_2)} \psi^{*}{}_{n_1,n_2}(\boldsymbol{r}_{\perp,{\rm out}}) \exp\left[-{\rm i}(n_1 + n_2 + 1)\sqrt{\frac{\Delta(z)}{f}}\right] \psi_{n_1,n_2}(\boldsymbol{r}_{\perp,{\rm in}}) \qquad (5.32)
\end{aligned}$$

where $\{\psi_{n_1,n_2} | n_1, n_2 = 1, 2, \cdots\}$ are the two-dimensional harmonic oscillator eigenfunctions. This expression gives the required propagator equivalent to (5.11) as can be verified directly by summing the above expression leading to the well-known propagator for the harmonic oscillator. The aberrations, in fact even the higher order corrections to the above result, say, in the case of thick lenses, can be treated as perturbations to the above oscillator-like Hamiltonian and using the standard techniques of the time-independent perturbation theory (replacing the ψ_{n_1,n_2}'s above by the perturbed eigenfunctions and the '$(n_1+n_2+1)\hbar\omega$'s by the corresponding eigenvalues) it should, in principle, be possible to obtain the required propagators for the aberrating systems. Thus the understanding of the performance of a charged-particle optical system is a straightforward affair in quantum theory. How such an understanding would help practical computation and design is to be seen. It is encouraging to see that it should be possible to use the Wigner function approach to deal with aberrating charged-particle optical systems also as has been demonstrated recently in [27].

The second remark we would like to add is regarding the extension of the Wigner function formalism to electron optics with spinor wavefunctions. To this end, we shall consider the formalism developed recently [102]. Then, to a Dirac spinor we can associate a 4×4 matrix Wigner function given by

$$W_{\alpha\beta}(\boldsymbol{r},\boldsymbol{p},t) = \frac{1}{(2\pi\hbar)^3}\int d^3\sigma\, \exp\left[-\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{\sigma}\right]\psi_\alpha\left(\boldsymbol{r}+\frac{\boldsymbol{\sigma}}{2},t\right)\psi_\beta^*\left(\boldsymbol{r}-\frac{\boldsymbol{\sigma}}{2},t\right),$$
$$\alpha,\beta = 1,2,3,4\cdot \qquad (5.33)$$

in the field-free region (i.e., $\mathbf{A} = (0,0,0)$). Since we are interested in relating the Wigner functions in field-free input and output regions of an electron optical system we adopt the above definition. For us, at any z-plane in the field-free regions the corresponding Wigner function matrix is

$$W_{\alpha\beta}(\boldsymbol{r}_\perp,\boldsymbol{p}_\perp,z) = \frac{1}{(2\pi\hbar)^2}\int d^2\sigma_\perp\, \exp\left[-\frac{\mathrm{i}}{\hbar}\boldsymbol{p}_\perp\cdot\boldsymbol{\sigma}_\perp\right]$$

$$\times\psi_\alpha\left(\boldsymbol{r}_\perp + \frac{\boldsymbol{\sigma}_\perp}{2}, z\right)\psi^*_\beta\left(\boldsymbol{r}_\perp - \frac{\boldsymbol{\sigma}_\perp}{2}, z\right),$$
$$\alpha, \beta = 1, 2, 3, 4. \tag{5.34}$$

In spinor electron optics the relation between the output and the input spinor wavefunctions is given by

$$\psi_\alpha\left(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}\right) = \int d^2 r_{\perp,\text{in}} \sum_{\beta=1}^{4} G_{\alpha\beta}\left(\boldsymbol{r}_{\perp,\text{out}}, z_{\text{out}}; \boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}\right)\psi_\beta\left(\boldsymbol{r}_{\perp,\text{in}}, z_{\text{in}}\right) \tag{5.35}$$

from which there should follow a relation between $\mathcal{W}_{\alpha\beta}(z_{\text{out}})$ and $\mathcal{W}_{\alpha\beta}(z_{\text{in}})$ generalizing (5.15). In this connection let us also note the existence of an alternative operator formalism [103] for the Wigner function which also has a natural gauge-covariant extension to the relativistic case; further, this formalism admits a straightforward second quantization leading directly to a manybody treatment. It should be interesting to explore the consequences of adopting this approach to the quantum theory of charged-particle optics in both the nonrelativistic and relativistic situations.

To summarize, we have studied the transformation properties of the Wigner function in the quantum theory of charged-particle optics in the scalar case, using the example of round magnetic lens under the paraxial approximation, and noted the possibility of the extension of the formalism to the case of spinor electron optics. It may also be noted that there exists a path integral approach to spinor electron optics [104].

Chapter 6

Concluding remarks

In this chapter we briefly summarize the contents of the thesis. We recall some significant points and lastly list some ideas/directions for future research in the quantum theory of charged-particle beam dynamics.

In Chapter I we briefly reviewed the classical theory of charged-particle beam optics, particularly the Lie algebraic formulation, pioneered by Dragt *et al.*, in the context of the charged-particle beam optics.

In Chapter II we have studied the quantum mechanics of charged-particle beam transport through optical systems with straight optic axis at the level of single particle. We have used an algebraic approach which molds the wave equation into a form suitable for treating quasimonoenergetic quasiparaxial beams propagating in the forward direction along the axis of the system. The example of the round magnetic lens has been done in detail including the aberrations. Besides we have also considered the electrostatic round lens, and magnetic and electrostatic quadrupole lenses in some detail. From the general theory presented and the examples considered we can draw several conclusions of which we would like to emphasize a few as follows.

Firstly, there are the explicit λ_0-dependent contributions to the paraxial behaviour and aberration coefficients (of all order aberrations), which have no analogues in the classical treatment. For instance, we worked out the explicit λ_0-dependent corrections to the spherical aberration. It is seen that the effects of such quantum corrections can be of any significance only at low energies.

Secondly, we find that *the aberrations depend not only on the quantum mechanical averages of* $\boldsymbol{r}_\perp$ *and* $\mathbf{p}_\perp$ *but also on their higher order central moments corresponding to the wave packets.* An immediate consequence of this fact is that contrary to the classical wisdom, coma, astigmatism, etc., *cannot* vanish for the object point situated on the optic axis.

To understand the above result as revealed by the quantum theory in contrast to the classical theory, let us examine the expressions for the observables in quantum mechanics. In quantum mechanics, for any observable O, $\left\langle\psi\left|f\left(\hat{O}\right)\right|\psi\right\rangle = f\left(\left\langle\psi\left|\hat{O}\right|\psi\right\rangle\right)$ only when the state $|\psi\rangle$ is an eigenstate of O and, in general for any two observables, say O_1 and O_2, only when the state $|\psi\rangle$ is a simultaneous eigenstate of both O_1 and O_2 can we have $\left\langle\psi\left|f\left(\hat{O}_1,\hat{O}_2\right)\right|\psi\right\rangle = f\left(\left\langle\psi\left|\hat{O}_1\right|\psi\right\rangle,\left\langle\psi\left|\hat{O}_2\right|\psi\right\rangle\right)$. This means that we *cannot* replace $\langle\hat{p}_x\hat{p}_\perp^2\rangle$, $\langle\{\hat{p}_x\,,\,\hat{\boldsymbol{p}}_\perp\cdot\boldsymbol{r}_\perp + \boldsymbol{r}_\perp\cdot\hat{\boldsymbol{p}}_\perp\}\rangle$, $\langle\{x\,,\,\hat{p}_\perp^2\}\rangle$, etc., respectively, by $\langle p_x\rangle\left(\langle p_x\rangle^2 + \langle p_y\rangle^2\right)$, $4\left(\langle x\rangle\langle p_x\rangle^2 + \langle y\rangle\langle p_x\rangle\langle p_y\rangle\right)$, $2\left(\langle x\rangle\left(\langle p_x\rangle^2 + \langle p_y\rangle^2\right)\right)$, etc. . As an illustration, consider the term $\sim \langle\{\boldsymbol{r}_\perp,\hat{p}_\perp^2\}\rangle(z_o)$, one of the terms contributing to coma which, being linear in position, is the dominant aberration next to the spherical aberration. The corresponding classical term, $\left(\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2\right)\boldsymbol{r}_\perp$ at z_o, vanishes obviously for an object-point on the axis. But, for a quantum wavepacket with $\langle\boldsymbol{r}_\perp\rangle(z_o) = (0,0)$ the value of $\langle\{\boldsymbol{r}_\perp,\hat{p}_\perp^2\}\rangle(z_o)$ need not be zero since it is not linear in $\langle\boldsymbol{r}_\perp\rangle(z_o)$. More explicitly, we can write, with $\delta\boldsymbol{r}_\perp = \boldsymbol{r}_\perp - \langle\boldsymbol{r}_\perp\rangle$ and $\delta\hat{\boldsymbol{p}}_\perp = \hat{\boldsymbol{p}}_\perp - \langle\boldsymbol{p}_\perp\rangle$,

$$
\begin{aligned}
\left\langle\left\{\boldsymbol{r}_\perp,\hat{p}_\perp^2\right\}\right\rangle(z_o) \;=\; & \left\langle\left\{\langle\boldsymbol{r}_\perp\rangle + \delta\boldsymbol{r}_\perp, (\langle p_x\rangle + \delta\hat{p}_x)^2\right.\right. \\
& \left.\left. +(\langle p_y\rangle + \delta\hat{p}_y)^2\right\}\right\rangle(z_o)
\end{aligned}
$$

$$
\begin{aligned}
&= \left\langle\left\{\langle \boldsymbol{r}_\perp\rangle + \delta\boldsymbol{r}_\perp, \langle p_x\rangle^2 + \langle p_y\rangle^2\right.\right. \\
&\qquad + (\delta\hat{p}_x)^2 + (\delta\hat{p}_y)^2 \\
&\qquad \left.\left. + 2\left(\langle p_x\rangle\delta\hat{p}_x + \langle p_y\rangle\delta\hat{p}_y\right)\right\}\right\rangle(z_o) \\
&= 2\langle \boldsymbol{r}_\perp\rangle(z_o)\langle \boldsymbol{p}_\perp\rangle(z_o)^2 \\
&\qquad + 2\langle \boldsymbol{r}_\perp\rangle(z_o)\left\langle(\delta\hat{p}_x)^2 + (\delta\hat{p}_y)^2\right\rangle(z_o) \\
&\qquad + \left\langle\left\{\delta\boldsymbol{r}_\perp, (\delta\hat{p}_x)^2 + (\delta\hat{p}_y)^2\right\}\right\rangle(z_o) \\
&\qquad + 2\left\langle\left\{\delta\boldsymbol{r}_\perp, \delta\hat{p}_x\right\}\right\rangle(z_o)\langle p_x\rangle(z_o) \\
&\qquad + 2\left\langle\left\{\delta\boldsymbol{r}_\perp, \delta\hat{p}_y\right\}\right\rangle(z_o)\langle p_y\rangle(z_o)\,, \qquad (6.1)
\end{aligned}
$$

showing clearly that this coma term is not necessarily zero for an object point on the axis, i.e., when $\langle \boldsymbol{r}_\perp\rangle(z_o) = (0,0)$. Equation 6.1 also shows how this coma term for off-axis points ($\langle \boldsymbol{r}_\perp\rangle(z_o) \neq (0,0)$) depends also on the higher order central moments besides the position ($\langle \boldsymbol{r}_\perp\rangle(z_o)$) and the slope ($\langle \boldsymbol{p}\rangle(z_o)/p_0$) of the corresponding classical ray. When an aperture is introduced in the path of the beam to limit the transverse momentum spread one will be introducing uncertainties in position coordinates ($\Delta x = \sqrt{\langle(\delta x)^2\rangle}$, $\Delta y = \sqrt{\langle(\delta y)^2\rangle}$) and hence the corresponding momentum uncertainties ($\Delta p_x = \sqrt{\langle(\delta\hat{p}_x)^2\rangle}$, $\Delta p_y = \sqrt{\langle(\delta\hat{p}_y)^2\rangle}$), in accordance with Heisenberg's uncertainty principle, and this would influence the aberrations.

It is to be seen how the above result would affect the scheme of correction of aberrations, though minutely. That is because the leading order λ_0-independent expressions for the aberration coefficients turn out to be same as the classical expressions, and the correction schemes depend on the matching of the aberration coefficients. So, the correction schemes are affected only by the λ_0-dependent terms, which are anyway very small compared to the leading order λ_0-independent terms.

That is for the scalar theory.

In Chapter III we discussed the spinor theory of charged-particle beam optics with the examples of the magnetic round lens and the magnetic quadrupole lens.

The main feature to be noted is that *the scalar approximation of the Dirac spinor theory differs from the Klein-Gordon theory, but the difference is only in the λ_0-dependent parts.* The leading λ_0-independent parts are identical. The other point to be noted is that the spinor contributions to the paraxial and aberration behaviour are proportional to the powers of λ_0. As an illustration, we have explicitly worked out the spherical-aberration-type contribution in the case of the magnetic round lens. We can expect such contributions to be relevant only in the case of applications like low energy electron microscopy (LEEM) where the electron energies are only in the range 1-100 eV [105]. Of course, in such nonrelativistic situations one can use our beam-optical version of the four-component Dirac formalism simply taking the two lower components of the Dirac spinor wavefunction to be zero.

In the fourth chapter, we have demonstrated how one can obtain a fully quantum mechanical understanding of the accelerator beam optics for a spin-$\frac{1}{2}$ particle, with anomalous magnetic moment, starting *ab initio* from the Dirac-Pauli equation. To this end, we have used a beam optical representation of the Dirac theory, and have shown that such an approach, in the lowest order approximation, leads naturally to a picture of orbital and spin dynamics based on the Lorentz force, the Stern-Gerlach force and the Thomas-BMT equation for spin evolution. It is further shown that even in the lowest order approximations of the theory there are deviations from the classical behaviour for spin evolution, namely, the dependence on differences $\langle xS_z\rangle - \langle x\rangle\langle S_z\rangle$, $\langle yS_z\rangle - \langle y\rangle\langle S_z\rangle$, $\langle p_xS_z\rangle - \langle p_x\rangle\langle S_z\rangle$, etc.. To illustrate the general theory we have considered the computation of the transfer maps for the spin components and the transverse phase-space, including the transverse Stern-Gerlach kicks, in the case of the normal and skew magnetic quadrupole lenses. It is found that our study supports the spin-splitter concepts based on transverse and longitudinal Stern-Gerlach kicks ([25], [74]-[80]). It would be worthwhile to extend the study initiated in the fourth chapter beyond the paraxial approximation to get higher order effects, which is sure to enrich our understanding of the polarized beam devices.

In the fifth chapter we presented an alternate approach to the quantum theory of charged-particle beam optics, by using the phase-space Wigner distribution function. The alternate formalism was demonstrated by working out the example of the magnetic round lens in the paraxial approximation. We also noted how the study can be extended to the spinor and aberrating case. It looks like that the formalism initiated in our study should be suitable for generalization in order to take into account the manybody effects also.

We suggest that the quantum theory of charged-particle beam dynamics may provide opportunities to test experimentally the various proposals for the choice of the position operator in relativistic quantum theory in view of the problem of localization [106] (for a discussion of this problem see, *e.g.*, [107]). This follows from the analysis in [71] where it has been found that different choices for the relativistic position operator, like following Newton-Wigner [106] or following Pryce [32], seem to give rise to different phase-space transfer maps for the Dirac particle beam passing through electromagnetic optical elements, in particular in the Stern-Gerlach sector.

Throughout the thesis we have confined our study to systems with straight optic axis. It would be interesting and practically useful to extend the study to systems with curved optic axis such as bending magnets, for example, which are essential components of charged-particle beam devices. Such a study would be useful for a better understanding of circular accelerators and storage rings. In these cases, the coordinate system used will have to be naturally the one adapted to the geometry, or the classical design orbit, of the system. Then, in the scalar theory one has to start with the Klein-Gordon equation written in the suitably chosen curvilinear coordinate system and the two-component form of the wavefunction will have to be introduced in such a way that one component describes the beam propagating in the forward direction along the curved optic axis and the other component describes the beam moving in the backward direction. Starting with such a two-component representation one can follow exactly the same approach, as used in the thesis, using the Foldy-Wouthuysen technique, to filter out the needed equation for the forward propagating beam. The rest of the analysis will follow the same scheme of calculations described in detail in the thesis. Similarly, for the Dirac theory we can start with the Dirac equation written using the chosen set of curvilinear coordinates following the method of construction of the Dirac equation in a generally covariant form (see, *e.g.*, [108]). Then, the treatment of the given system follows in the same way, via the Foldy-Wouthuysen transformations, as discussed earlier. The work in this direction initiated by Jagannathan [8] needs to be carried out to generalize the present study to cover the systems with curved optic axis.

It may also be pointed out that a deeper understanding of electron optics from the point of view quantum mechanics should be necessary to tackle the problems of coherence in applications like electron holography.

Throughout the thesis we have also restricted ourselves to the treatment of propagation of a monoenergetic paraxial beam through a single optical block with a static magnetic field and straight axis. Thus, it is obvious that there are several open problems related to the issues concerning the extension of the present formalism to more complicated situations. Leaving aside the problems of including the effects of multiparticle dynamics, quantum nature of the electromagnetic field, interaction with radiation, etc., for the present, the immediate concern should be about the extension of the formalism taking into account the chromatic effects, curvature of the optic axis, global analysis of systems like storage rings, and time-dependence of fields.

When the beam entering the time-independent system from the field-free input region is not monochromatic, as is in general, the wavefunction of the beam propagating through the system in the $+z$-direction can be written, in the Dirac representation, as

$$\Psi_D(\boldsymbol{r},t) = \int_{p_0-\frac{1}{2}\Delta p}^{p_0+\frac{1}{2}\Delta p} dp\, \psi_D(\boldsymbol{r};p) \exp\left(-\mathrm{i}E(p)t/\hbar\right) , \quad \Delta p \ll p_0 , \tag{6.2}$$

where p_0 is the design momentum and

$$\psi_D(\boldsymbol{r}_\perp, z < z_n; p) = \frac{1}{(2\pi\hbar)^{3/2}} \int\int dp_x\, dp_y\, \psi_{FD}(\boldsymbol{r}_\perp, z; \boldsymbol{p}) , \quad |\boldsymbol{p}_\perp| \ll p , \tag{6.3}$$

with $\psi_{FD}(\boldsymbol{r}_\perp, z; p)$ obtained from (4.13) by replacing the constants (s_+, s_-) by the functions $(s_+(\boldsymbol{p}), s_-(\boldsymbol{p}))$. Now, the z-evolution of each Fourier component $(\psi_D(\boldsymbol{r};p))$ of $\Psi_D(\boldsymbol{r},t)$ will have to be traced according the above formalism for monochromatic beam and the results will have to be integrated to get the z-evolution of the time-dependent $\Psi_D(\boldsymbol{r},t)$; generalization is straightforward in the case of description using density matrices. Using such a procedure it should be possible to account for the chromatic effects of static optical elements. First, one should be able to derive in this way the well known classical results on chromatic effects (see, *e.g.*, [1–3, 5]) in the lowest order approximation. Note that in the monoenergetic case, with $\Delta p = 0$, the phase factor $\exp\left(-\mathrm{i}E(p_0)t/\hbar\right)$ drops out of the formalism making time simply spectator.

Analysis of global systems, like storage rings, should be possible by learning to patch together the quantum transfer maps for individual, or local, optical blocks to produce the quantum one-turn map (see [20] and references therein for help from classical beam dynamics).

Finally, the question of time-dependent fields : The present formalism can lead only to a relationship among the wavefunctions at transverse planes situated along the design orbit guided by static fields. To take into account time-dependent effects, radiation, etc., one will have to use only the traditional quantum dynamical time evolution equation. The present formalism is mainly intended to study effectively the static optical characteristics of the system. We believe that a hybrid approach to beam dynamics obtained by integrating the present formalism, suited for static characteristics of beam optics, with the traditional methods of quantum dynamics for studying the time-dependent aspects should be profitable.

Appendix A

The Magnus Formula

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

$$\mathrm{e}^{\hat{A}}\mathrm{e}^{\hat{B}} = \mathrm{e}^{\hat{A}+\hat{B}+\frac{1}{2}[\hat{A},\hat{B}]+\frac{1}{12}\left\{[[\hat{A},\hat{A}],\hat{B}]+[[\hat{A},\hat{B}],\hat{B}]\right\}+\cdots}. \tag{A1}$$

Let it be required to solve the differential equation

$$\frac{\partial}{\partial t}u(t) = \hat{A}(t)u(t) \tag{A2}$$

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

$$u(t_0 + \Delta t) = \mathrm{e}^{\Delta t \hat{A}(t_0)}u(t_0). \tag{A3}$$

Iterating this solution we have

$$\begin{aligned} u(t_0 + 2\Delta t) &= \mathrm{e}^{\Delta t \hat{A}(t_0+\Delta t)}\mathrm{e}^{\Delta t \hat{A}(t_0)}u(t_0) \\ u(t_0 + 3\Delta t) &= \mathrm{e}^{\Delta t \hat{A}(t_0+2\Delta t)}\mathrm{e}^{\Delta t \hat{A}(t_0+\Delta t)}\mathrm{e}^{\Delta t \hat{A}(t_0)}u(t_0) \\ &\ldots \quad \text{and so on.} \end{aligned} \tag{A4}$$

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

$$u(T) = \left\{\prod_{n=0}^{N-1} e^{\Delta t \hat{A}(t_0+n\Delta t)}\right\} u(t_0). \tag{A5}$$

Thus, $u(T)$ is given by computing the product in (A5) using successively the BCH-formula (A1)) 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, [109]) :

$$\begin{aligned}
u(T) &= \hat{\mathcal{T}}(T, t_0)u(t_0) \\
\mathcal{T}(T, t_0) &= \exp\left\{\int_{t_0}^{T} dt_1\, \hat{A}(t_1)\right. \\
&\quad + \frac{1}{2}\int_{t_0}^{T} dt_2 \int_{t_0}^{t_2} dt_1\, \left[\hat{A}(t_2), \hat{A}(t_1)\right] \\
&\quad + \frac{1}{6}\int_{t_0}^{T} dt_3 \int_{t_0}^{t_3} dt_2 \int_{t_0}^{t_2} dt_1\, \left(\left[\left[\hat{A}(t_3), \hat{A}(t_2)\right], \hat{A}(t_1)\right]\right. \\
&\quad \left.\left. + \left[\left[\hat{A}(t_1), \hat{A}(t_2)\right], \hat{A}(t_3)\right]\right) + \ldots\right\} . \qquad \text{(A6)}
\end{aligned}$$

To see how the equation (A6) is obtained let us substitute the assumed form of the solution, $u(t) = \hat{\mathcal{T}}(t, t_0)\, u(t_0)$, in (A2). Then, it is seen that $\hat{\mathcal{T}}(t, t_0)$ obeys the equation

$$\frac{\partial}{\partial t}\hat{\mathcal{T}}(t, t_0) = \hat{A}(t)\mathcal{T}(t, t_0), \qquad \hat{\mathcal{T}}(t_0, t_0) = \hat{I}\,. \qquad \text{(A7)}$$

Introducing an iteration parameter λ in (A7), let

$$\frac{\partial}{\partial t}\hat{\mathcal{T}}(t, t_0; \lambda) = \lambda\hat{A}(t)\hat{\mathcal{T}}(t, t_0; \lambda)\,, \qquad \text{(A8)}$$

$$\hat{\mathcal{T}}(t_0, t_0; \lambda) = \hat{I}, \quad \hat{\mathcal{T}}(t, t_0; 1) = \hat{\mathcal{T}}(t, t_0)\,. \qquad \text{(A9)}$$

Assume a solution of (A8) to be of the form

$$\hat{\mathcal{T}}(t, t_0; \lambda) = e^{\Omega(t, t_0; \lambda)} \qquad \text{(A10)}$$

with

$$\Omega(t, t_0; \lambda) = \sum_{n=1}^{\infty} \lambda^n \Delta_n(t, t_0), \quad \Delta_n(t_0, t_0) = 0 \quad \text{for all } n\,. \qquad \text{(A11)}$$

Now, using the identity (se Wilcox, [110])

$$\frac{\partial}{\partial t}e^{\Omega(t, t_0; \lambda)} = \left\{\int_0^1 ds e^{s\Omega(t, t_0; \lambda)} \frac{\partial}{\partial t}\Omega(t, t_0; \lambda) e^{-s\Omega(t, t_0; \lambda)}\right\} e^{\Omega(t, \lambda)}, \qquad \text{(A12)}$$

one has

$$\int_0^1 ds e^{s\Omega(t,t_0;\lambda)} \frac{\partial}{\partial t}\Omega(t,t_0;\lambda) e^{-s\Omega(t,t_0;\lambda)} = \lambda \hat{A}(t) . \tag{A13}$$

Substituting in (A13) the series expression for $\Omega(t, t_0; \lambda)$ (A11), expanding the left hand side using the first identity in (C8), integrating and equating the coefficients 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 $j = 1$

$$\frac{\partial}{\partial t}\Delta_1(t,t_0) = \hat{A}(t), \qquad \Delta_1(t_0,t_0) = 0 \tag{A14}$$

and hence

$$\Delta_1(t,t_0) = \int_{t_0}^{t} dt_1 \hat{A}(t_1) . \tag{A15}$$

For $j = 2$

$$\frac{\partial}{\partial t}\Delta_2(t,t_0) + \frac{1}{2}\left[\Delta_1(t,t_0) , \frac{\partial}{\partial t}\Delta_1(t,t_0)\right] = 0 , \qquad \Delta_2(t_0,t_0) = 0 \tag{A16}$$

and hence

$$\Delta_2(t,t_0) = \frac{1}{2}\int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1 \left[\hat{A}(t_2) , \hat{A}(t_1)\right] . \tag{A17}$$

Similarly,

$$\begin{aligned} \Delta_3(t,t_0) &= \frac{1}{6}\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 \left\{\left[\left[\hat{A}(t_1) , \hat{A}(t_2)\right] , \hat{A}(t_3)\right]\right. \\ &\qquad \left. + \left[\left[\hat{A}(t_3) , \hat{A}(t_2)\right] , \hat{A}(t_1)\right]\right\} . \end{aligned} \tag{A18}$$

Then, the Magnus formula in (A6) follows from (A9)-(A11). Equation 2.69 we have, in the context of z-evolution follows from the above discussion with the identification $t \longrightarrow z$, $t_0 \longrightarrow z^{(1)}$, $T \longrightarrow z^{(2)}$ and $\hat{A}(t) \longrightarrow -\frac{i}{\hbar}\hat{\mathcal{H}}_o(z)$.

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 [110], Bellman and Vasudevan [111], Dattoli *et al.* [39], and references therein.

Appendix B

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, [29]) to the Schrö-dinger 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, [31]).

Defining

$$\Phi = \frac{\partial}{\partial t}\Psi \,, \tag{B1}$$

the free particle Klein-Gordon equation is written as

$$\frac{\partial}{\partial t}\Phi = \left(c^2\nabla^2 - \frac{{m_0}^2c^4}{\hbar^2}\right)\Psi \,. \tag{B2}$$

Introducing the linear combinations

$$\Psi_+ = \frac{1}{2}\left(\Psi + \frac{\mathrm{i}\hbar}{m_0c^2}\Phi\right) , \qquad \Psi_- = \frac{1}{2}\left(\Psi - \frac{\mathrm{i}\hbar}{m_0c^2}\Phi\right) \tag{B3}$$

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

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi_+ \quad = \quad -\frac{\hbar^2\nabla^2}{2m_0}\left(\Psi_+ + \Psi_-\right) + m_0c^2\Psi_+$$

$$i\hbar\frac{\partial}{\partial t}\Psi_- = \frac{\hbar^2\nabla^2}{2m_0}(\Psi_+ + \Psi_-) - m_0c^2\Psi_- . \tag{B4}$$

Equation (B4) can be written in a two-component language as

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} = \hat{H}_0^{FV}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix}, \tag{B5}$$

with the Feshbach-Villars Hamiltonian for the free particle, $\hat{H}_0^{FV}$, given by

$$\begin{aligned}\hat{H}_0^{FV} &= \begin{pmatrix} m_0c^2 + \frac{\hat{p}^2}{2m_0} & \frac{\hat{p}^2}{2m_0} \\ -\frac{\hat{p}^2}{2m_0} & -m_0c^2 - \frac{\hat{p}^2}{2m_0}\end{pmatrix} \\ &= m_0c^2\sigma_z + \frac{\hat{p}^2}{2m_0}\sigma_z + i\frac{\hat{p}^2}{2m_0}\sigma_y .\end{aligned} \tag{B6}$$

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

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

$$\hat{\boldsymbol{p}} \longrightarrow \hat{\boldsymbol{\pi}} = \hat{\boldsymbol{p}} - q\boldsymbol{A}, \qquad i\hbar\frac{\partial}{\partial t} \longrightarrow i\hbar\frac{\partial}{\partial t} - q\phi. \tag{B7}$$

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

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} &= \hat{H}^{FV}\begin{pmatrix}\Psi_+\\ \Psi_-\end{pmatrix} \\ \begin{pmatrix}\Psi_+\\ \\ \Psi_-\end{pmatrix} &= \frac{1}{2}\begin{pmatrix}\Psi + \frac{1}{m_0c^2}\left(i\hbar\frac{\partial}{\partial t} - q\phi\right)\Psi \\ \\ \Psi - \frac{1}{m_0c^2}\left(i\hbar\frac{\partial}{\partial t} - q\phi\right)\Psi\end{pmatrix} \\ \hat{H}^{FV} &= m_0c^2\sigma_z + \hat{\mathcal{E}} + \hat{\mathcal{O}} \\ \hat{\mathcal{E}} &= q\phi + \frac{\hat{\pi}^2}{2m_0}\sigma_z, \quad \hat{\mathcal{O}} = i\frac{\hat{\pi}^2}{2m_0}\sigma_y. \end{aligned} \tag{B8}$$

As in the free-particle case, in the nonrelativistic situation Ψ_+ is large compared to Ψ_-. The even term $\hat{\mathcal{E}}$ does not couple Ψ_+ and Ψ_- whereas $\hat{\mathcal{O}}$ is odd which couples Ψ_+ and Ψ_-. Starting from (B8), 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, [31]).

It is clear from the above that we have just adopted the above technique for studying the z-evolution of the Klein-Gordon wavefunction 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 C

The Foldy-Wouthuysen Representation of the Dirac Equation

The main framework of the formalism of charged-particle wave optics, used here for both the scalar theory and the spinor theory, 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 and an applied electromagnetic field in a nonrelativistic and easily interpretable form (Foldy and Wouthuysen, [30]; see also Pryce, [32], Tani, [33]; see Acharya and Sudarshan, [34], for a general discussion of the role of 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. Analogously, in the optical formalism the aim has been to filter out from the nonrelativistic Schrödinger equation, or the Klein-Gordon equation, or the Dirac equation, the part which describes the evolution of the charged-particle beam along the axis of an optical system comprising a stationary electromagnetic field, using the Foldy-Wouthuysen technique.

Let us describe here briefly the standard Foldy-Wouthuysen theory so that the way it has been adopted for the purposes of the above studies in charged-particle wave optics will be clear. The Dirac equation in presence of an electromagnetic field is

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) &= \hat{H}_D\Psi(\boldsymbol{r},t) && \text{(C1)}\\
\hat{H}_D &= m_0c^2\beta + q\phi + c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}} \\
&= m_0c^2\beta + \hat{\mathcal{E}} + \hat{\mathcal{O}}\,, && \text{(C2)}
\end{aligned}
$$

with $\hat{\mathcal{E}} = q\phi$ and $\hat{\mathcal{O}} = c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}}$. 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 $\hat{\mathcal{E}}$ which does not couple the large and small components of Ψ is called 'even' and $\hat{\mathcal{O}}$ is called an 'odd' operator which couples the large to the small components. Note that

$$
\beta\hat{\mathcal{O}} = -\hat{\mathcal{O}}\beta\,, \qquad \beta\hat{\mathcal{E}} = \hat{\mathcal{E}}\beta\,. \tag{C3}
$$

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

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

$$
\begin{aligned}
&\Psi \longrightarrow \Psi' = \hat{U}_F\Psi \\
&\qquad \hat{U}_F = e^{i\hat{S}} = e^{\beta\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}\theta}\,, \quad \tan 2|\hat{\boldsymbol{p}}|\theta = \frac{|\hat{\boldsymbol{p}}|}{m_0c}\,. && \text{(C4)}
\end{aligned}
$$

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

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\Psi' &= e^{i\hat{S}}\left(m_0c^2\beta + c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}\right)e^{-i\hat{S}}\Psi' \\
&= \left(\cos|\hat{\boldsymbol{p}}|\theta + \frac{\beta\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}}{|\hat{\boldsymbol{p}}|}\sin|\hat{\boldsymbol{p}}|\theta\right)\left(m_0c^2\beta + c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}\right) \\
&\qquad \times\left(\cos|\hat{\boldsymbol{p}}|\theta - \frac{\beta\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}}{|\hat{\boldsymbol{p}}|}\sin|\hat{\boldsymbol{p}}|\theta\right)\Psi'
\end{aligned}
$$

$$
\begin{aligned}
&= \left(m_0c^2\cos 2|\hat{p}|\theta + c|\hat{p}|\sin 2|\hat{p}|\theta\right)\beta\Psi' \\
&= \left(\sqrt{m_0^2c^4 + c^2\hat{p}^2}\right)\beta\,\Psi' . \qquad \text{(C5)}
\end{aligned}
$$

In the general case, when the electron is in a time-dependent electromagnetic field it is not possible to construct an $\exp(\mathrm{i}\hat{S})$ 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 $|\boldsymbol{p}| \ll m_0c$, the transformation operator $\hat{U}_F = \exp(\mathrm{i}\hat{S})$ with $\hat{S} \approx -\mathrm{i}\beta\hat{\mathcal{O}}/2m_0c^2$, where $\hat{\mathcal{O}} = c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}}$ is the odd part of the free Hamiltonian. So, in the general case we can start with the transformation

$$
\Psi^{(1)} = \mathrm{e}^{\mathrm{i}\hat{S}_1}\Psi, \qquad \hat{S}_1 = -\frac{\mathrm{i}\beta\hat{\mathcal{O}}}{2m_0c^2} = -\frac{\mathrm{i}\beta\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}}}{2m_0c} . \qquad \text{(C6)}
$$

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

$$
\begin{aligned}
\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi^{(1)} &= \mathrm{i}\hbar\frac{\partial}{\partial t}\left(\mathrm{e}^{\mathrm{i}\hat{S}_1}\Psi\right) = \mathrm{i}\hbar\frac{\partial}{\partial t}\left(\mathrm{e}^{\mathrm{i}\hat{S}_1}\right)\Psi + \mathrm{e}^{\mathrm{i}\hat{S}_1}\left(\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi\right) \\
&= \left[\mathrm{i}\hbar\frac{\partial}{\partial t}\left(\mathrm{e}^{\mathrm{i}\hat{S}_1}\right) + \mathrm{e}^{\mathrm{i}\hat{S}_1}\hat{H}_D\right]\Psi \\
&= \left[\mathrm{i}\hbar\frac{\partial}{\partial t}\left(\mathrm{e}^{\mathrm{i}\hat{S}_1}\right)\mathrm{e}^{-\mathrm{i}\hat{S}_1} + \mathrm{e}^{\mathrm{i}\hat{S}_1}\hat{H}_D\mathrm{e}^{-\mathrm{i}\hat{S}_1}\right]\Psi^{(1)} \\
&= \left[\mathrm{e}^{\mathrm{i}\hat{S}_1}\hat{H}_D\mathrm{e}^{-\mathrm{i}\hat{S}_1} - \mathrm{i}\hbar\mathrm{e}^{\mathrm{i}\hat{S}_1}\frac{\partial}{\partial t}\left(\mathrm{e}^{-\mathrm{i}\hat{S}_1}\right)\right]\Psi^{(1)} \\
&= \hat{H}_D^{(1)}\Psi^{(1)} \qquad \text{(C7)}
\end{aligned}
$$

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

Now, using the identities

$$
\begin{aligned}
\mathrm{e}^{\hat{A}}\hat{B}\mathrm{e}^{-\hat{A}} &= \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]] + \frac{1}{3!}[\hat{A},[\hat{A},[\hat{A},\hat{B}]]] + \ldots \\
&\mathrm{e}^{\hat{A}(t)}\frac{\partial}{\partial t}\left(\mathrm{e}^{-\hat{A}(t)}\right) \\
&= \left(1 + \hat{A}(t) + \frac{1}{2!}\hat{A}(t)^2 + \frac{1}{3!}\hat{A}(t)^3\cdots\right)
\end{aligned}
$$

$$
\begin{aligned}
&\times\frac{\partial}{\partial t}\left(1-\hat{A}(t)+\frac{1}{2!}\hat{A}(t)^2-\frac{1}{3!}\hat{A}(t)^3\cdots\right)\\
&=\left(1+\hat{A}(t)+\frac{1}{2!}\hat{A}(t)^2+\frac{1}{3!}\hat{A}(t)^3\cdots\right)\\
&\quad\times\left(-\frac{\partial\hat{A}(t)}{\partial t}+\frac{1}{2!}\left\{\frac{\partial\hat{A}(t)}{\partial t}\hat{A}(t)+\hat{A}(t)\frac{\partial\hat{A}(t)}{\partial t}\right\}\right.\\
&\quad-\frac{1}{3!}\left\{\frac{\partial\hat{A}(t)}{\partial t}\hat{A}(t)^2+\hat{A}(t)\frac{\partial\hat{A}(t)}{\partial t}\hat{A}(t)\right.\\
&\quad\left.\left.+\hat{A}(t)^2\frac{\partial\hat{A}(t)}{\partial t}\right\}\cdots\right)\\
&\approx-\frac{\partial\hat{A}(t)}{\partial t}-\frac{1}{2!}\left[\hat{A}(t),\frac{\partial\hat{A}(t)}{\partial t}\right]\\
&\quad-\frac{1}{3!}\left[\hat{A}(t),\left[\hat{A}(t),\frac{\partial\hat{A}(t)}{\partial t}\right]\right]\\
&\quad-\frac{1}{4!}\left[\hat{A}(t),\left[\hat{A}(t),\left[\hat{A}(t),\frac{\partial\hat{A}(t)}{\partial t}\right]\right]\right],
\end{aligned}
\tag{C8}
$$

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

$$
\begin{aligned}
\hat{H}_D^{(1)} &\approx \hat{H}_D-\hbar\frac{\partial\hat{S}_1}{\partial t}+\mathrm{i}\left[\hat{S}_1,\hat{H}_D-\frac{\hbar}{2}\frac{\partial\hat{S}_1}{\partial t}\right]\\
&\quad-\frac{1}{2!}\left[\hat{S}_1,\left[\hat{S}_1,\hat{H}_D-\frac{\hbar}{3}\frac{\partial\hat{S}_1}{\partial t}\right]\right]\\
&\quad-\frac{\mathrm{i}}{3!}\left[\hat{S}_1,\left[\hat{S}_1,\left[\hat{S}_1,\hat{H}_D-\frac{\hbar}{4}\frac{\partial\hat{S}_1}{\partial t}\right]\right]\right].
\end{aligned}
\tag{C9}
$$

Substituting in (C9), $\hat{H}_D=m_0c^2\beta+\hat{\mathcal{E}}+\hat{\mathcal{O}}$, simplifying the right hand side using the relations $\beta\hat{\mathcal{O}}=-\hat{\mathcal{O}}\beta$ and $\beta\hat{\mathcal{E}}=\hat{\mathcal{E}}\beta$ and collecting everything together, we have

$$
\begin{aligned}
\hat{H}_D^{(1)} &\approx m_0c^2\beta+\hat{\mathcal{E}}_1+\hat{\mathcal{O}}_1\\
\hat{\mathcal{E}}_1 &\approx \hat{\mathcal{E}}+\frac{1}{2m_0c^2}\beta\hat{\mathcal{O}}^2-\frac{1}{8m_0^2c^4}\left[\hat{\mathcal{O}},\left(\left[\hat{\mathcal{O}},\hat{\mathcal{E}}\right]+\mathrm{i}\hbar\frac{\partial\hat{\mathcal{O}}}{\partial t}\right)\right]\\
&\quad-\frac{1}{8m_0^3c^6}\beta\hat{\mathcal{O}}^4\\
\hat{\mathcal{O}}_1 &\approx \frac{\beta}{2m_0c^2}\left(\left[\hat{\mathcal{O}},\hat{\mathcal{E}}\right]+\mathrm{i}\hbar\frac{\partial\hat{\mathcal{O}}}{\partial t}\right)-\frac{1}{3m_0^2c^4}\hat{\mathcal{O}}^3,
\end{aligned}
\tag{C10}
$$

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

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

$$\begin{aligned}
\Psi^{(2)} &= e^{i\hat{S}_2}\Psi^{(1)}, \\
\hat{S}_2 &= -\frac{i\beta\hat{\mathcal{O}}_1}{2m_0c^2} \\
&= -\frac{i\beta}{2m_0c^2}\left[\frac{\beta}{2m_0c^2}\left(\left[\hat{\mathcal{O}},\hat{\mathcal{E}}\right] + i\hbar\frac{\partial\hat{\mathcal{O}}}{\partial t}\right) - \frac{1}{3m_0^2c^4}\hat{\mathcal{O}}^3\right].
\end{aligned} \tag{C11}$$

After this transformation,

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}\Psi^{(2)} &= \hat{H}_D^{(2)}\Psi^{(2)}, \quad \hat{H}_D^{(2)} = m_0c^2\beta + \hat{\mathcal{E}}_2 + \hat{\mathcal{O}}_2 \\
\hat{\mathcal{E}}_2 &\approx \hat{\mathcal{E}}_1, \quad \hat{\mathcal{O}}_2 \approx \frac{\beta}{2m_0c^2}\left(\left[\hat{\mathcal{O}}_1,\hat{\mathcal{E}}_1\right] + i\hbar\frac{\partial\hat{\mathcal{O}}_1}{\partial t}\right),
\end{aligned} \tag{C12}$$

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

$$\Psi^{(3)} = e^{i\hat{S}_3}\Psi^{(2)}, \qquad \hat{S}_3 = -\frac{i\beta\hat{\mathcal{O}}_2}{2m_0c^2} \tag{C13}$$

we have

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}\Psi^{(3)} &= \hat{H}_D^{(3)}\Psi^{(3)}, \quad \hat{H}_D^{(3)} = m_0c^2\beta + \hat{\mathcal{E}}_3 + \hat{\mathcal{O}}_3 \\
\hat{\mathcal{E}}_3 &\approx \hat{\mathcal{E}}_2 \approx \hat{\mathcal{E}}_1, \quad \hat{\mathcal{O}}_3 \approx \frac{\beta}{2m_0c^2}\left(\left[\hat{\mathcal{O}}_2,\hat{\mathcal{E}}_2\right] + i\hbar\frac{\partial\hat{\mathcal{O}}_2}{\partial t}\right),
\end{aligned} \tag{C14}$$

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

$$\hat{H}_D^{(3)} \approx m_0c^2\beta + \hat{\mathcal{E}} + \frac{1}{2m_0c^2}\beta\hat{\mathcal{O}}^2$$

$$
\begin{aligned}
& -\frac{1}{8m_0^2c^4}\left[\hat{\mathcal{O}},\left(\left[\hat{\mathcal{O}},\hat{\mathcal{E}}\right]+\mathrm{i}\hbar\frac{\partial\hat{\mathcal{O}}}{\partial t}\right)\right] \\
& -\frac{1}{8m_0^3c^6}\beta\hat{\mathcal{O}}^4 . \qquad \text{(C15)}
\end{aligned}
$$

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

$$
\begin{aligned}
\hat{\mathcal{E}}_j &= \hat{\mathcal{E}}_1\left(\hat{\mathcal{E}}\to\hat{\mathcal{E}}_{j-1},\hat{\mathcal{O}}\to\hat{\mathcal{O}}_{j-1}\right) \\
\hat{\mathcal{O}}_j &= \hat{\mathcal{O}}_1\left(\hat{\mathcal{E}}\to\hat{\mathcal{E}}_{j-1},\hat{\mathcal{O}}\to\hat{\mathcal{O}}_{j-1}\right), \quad j>1, \qquad \text{(C16)}
\end{aligned}
$$

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

With $\hat{\mathcal{E}} = q\phi$ and $\hat{\mathcal{O}} = c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\pi}}$, the final reduced Hamiltonian (C15) is, to the order calculated,

$$
\begin{aligned}
\hat{H}_D^{(3)} &= \beta\left(m_0c^2+\frac{\hat{\pi}^2}{2m_0}-\frac{\hat{p}^4}{8m_0^3c^6}\right)+q\phi-\frac{q\hbar}{2m_0c}\beta\boldsymbol{\Sigma}\cdot\boldsymbol{B} \\
&\quad -\frac{\mathrm{i}q\hbar^2}{8m_0^2c^2}\boldsymbol{\Sigma}\cdot\mathrm{curl}\,\boldsymbol{E}-\frac{q\hbar}{4m_0^2c^2}\boldsymbol{\Sigma}\cdot\boldsymbol{E}\times\hat{\boldsymbol{p}} \\
&\quad -\frac{q\hbar^2}{8m_0^2c^2}\mathrm{div}\boldsymbol{E}, \qquad \text{(C17)}
\end{aligned}
$$

with the individual terms having direct physical interpretations. The terms in the first parenthesis result from the expansion of $\sqrt{m_0^2c^4+c^2\hat{\pi}^2}$ 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. Noting the analogy between the nonrelativistic particle dynamics and paraxial optics, 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 relevant wave equation in a beam optical form resembling exactly the Dirac equation ((C1)-(C2)) 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 ($\hat{\mathcal{E}}$), an odd part ($\hat{\mathcal{O}}$), a suitable expansion parameter characterizing the dominant forward propagation and a leading term with a β-like coefficient commuting with $\hat{\mathcal{E}}$ and anticommuting with $\hat{\mathcal{O}}$). The additional feature of our formalism is to return finally to the original representation after making an extra approximation, dropping β from the final reduced optical Hamiltonian, taking into account the fact that we are interested only in the forward-propagating beam.

Appendix D

Green's Function for the Nonrelativistic Free Particle

For a nonrelativistic free particle of mass m moving in one dimension the Schrödinger equation is

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(x,t) = \hat{H}\Psi(x,t), \qquad \hat{H} = \frac{\hat{p}_x^2}{2m}. \tag{D1}$$

The corresponding Green's function, or the propagator, given by

$$G(x',t';x,t) = \langle x'|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(t'-t)\hat{H}}|x\rangle = \langle x'|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(t'-t)\frac{\hat{p}_x^2}{2m}}|x\rangle \tag{D2}$$

is such that

$$\Psi(x',t') = \int dx\, G(x',t';x,t)\Psi(x,t). \tag{D3}$$

The expression for $G(x',t';x,t)$ in (D2) can be evaluated easily using the momentum representation. The explicit calculation is as follows: with $(t'-t) = \Delta t$,

$$\begin{aligned}
&\langle x'|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\Delta t\frac{\hat{p}_x^2}{2m}}|x\rangle \\
&= \int\!\!\int dp'dp\,\langle x'|p'\rangle\langle p'|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\Delta t\frac{\hat{p}_x^2}{2m}}|p\rangle\langle p|x\rangle \\
&= \int\!\!\int dp'dp\left(\frac{\mathrm{e}^{\frac{\mathrm{i}}{\hbar}p'x'}}{\sqrt{2\pi\hbar}}\right)\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\Delta t\frac{p'^2}{2m}}\delta(p'-p)\left(\frac{\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}px}}{\sqrt{2\pi\hbar}}\right) \\
&= \frac{1}{2\pi\hbar}\int dp'\exp\left\{\frac{\mathrm{i}}{\hbar}\left(p'(x'-x) - \Delta t\frac{p'^2}{2m}\right)\right\} \\
&= \frac{1}{2\pi\hbar}\int dp'\exp\left\{-\frac{\mathrm{i}\Delta t}{2m\hbar}\left[\left(p' - \frac{m(x'-x)}{\Delta t}\right)^2\right.\right.
\end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad\qquad\qquad - \frac{m^2(x'-x)^2}{(\Delta t)^2}\Bigg]\Bigg\} \\
&= \frac{1}{2\pi\hbar}\exp\left\{\frac{\mathrm{i}m(x'-x)^2}{2\hbar\Delta t}\right\} \\
&\qquad \times \int dp' \exp\left\{-\frac{\mathrm{i}\Delta t}{2m\hbar}\left(p' - \frac{m(x'-x)}{\Delta t}\right)^2\right\} \\
&= \frac{1}{2\pi\hbar}\exp\left\{\frac{\mathrm{i}m(x'-x)^2}{2\hbar\Delta t}\right\}\int dP' \exp\left\{-\frac{\mathrm{i}\Delta t}{2m\hbar}P'^2\right\} \\
&= \left(\frac{m}{2\pi\mathrm{i}\hbar\Delta t}\right)^{\frac{1}{2}}\exp\left\{\frac{\mathrm{i}m(x'-x)^2}{2\hbar\Delta t}\right\}
\end{aligned}
\tag{D4}
$$

We have used the two-dimensional generalization of this result. Since the variables x and y are separable for the free motion in the xy-plane, it follows that

$$
G(\boldsymbol{r}'_\perp, t'; \boldsymbol{r}_\perp, t) = \left(\frac{m}{2\pi\mathrm{i}\hbar\Delta t}\right)\exp\left\{\frac{\mathrm{i}m|\boldsymbol{r}'_\perp - \boldsymbol{r}_\perp|^2}{2\hbar\Delta t}\right\}. \tag{D5}
$$

Appendix E

Matrix element of the Rotation Operator

The required matrix element of the rotation operator around the z-axis through an angle ϑ, $\langle \boldsymbol{r}_\perp | e^{\frac{i}{\hbar}\vartheta \hat{L}_z} | \tilde{\boldsymbol{r}}_\perp \rangle$, can be easily calculated using the cylindrical coordinate system $x = \rho\cos\theta$, $y = \rho\sin\theta$, $z = z$. Then,

$$\begin{aligned} e^{\frac{i}{\hbar}\vartheta \hat{L}_z} \psi(\boldsymbol{r}_\perp, z) &= e^{\frac{i}{\hbar}\vartheta(x\hat{p}_y - y\hat{p}_x)} \psi(\boldsymbol{r}_\perp, z) = e^{\vartheta \frac{\partial}{\partial\theta}} \psi(\rho, \theta, z) \\ &= \psi(\rho, \theta + \vartheta, z) = \psi(\boldsymbol{r}_\perp(\vartheta), z)\,, \end{aligned} \tag{E1}$$

where $\boldsymbol{r}_\perp(\vartheta) = (x\cos\vartheta - y\sin\vartheta, x\sin\vartheta + y\cos\vartheta)$. Using this result, we get

$$\begin{aligned} \langle \boldsymbol{r}'_\perp | e^{\frac{i}{\hbar}\vartheta \hat{L}_z} | \tilde{\boldsymbol{r}}_\perp \rangle &= \int d^2 r \delta^{2*}(\boldsymbol{r}_\perp - \boldsymbol{r}'_\perp) e^{\frac{i}{\hbar}\vartheta \hat{L}_z} \delta^2(\boldsymbol{r}_\perp - \tilde{\boldsymbol{r}}_\perp) \\ &= \int d^2 r \delta^{2*}(\boldsymbol{r}_\perp - \boldsymbol{r}'_\perp) \delta^2\left(\boldsymbol{r}_\perp(\vartheta) - \tilde{\boldsymbol{r}}_\perp\right) \\ &= \delta^2\left(\boldsymbol{r}'_\perp(\vartheta) - \tilde{\boldsymbol{r}}_\perp\right)\,. \end{aligned} \tag{E2}$$

Hence we get 2.116:

$$\langle \boldsymbol{r}'_\perp | e^{\frac{i}{\hbar}\theta(z,z_0)\hat{L}_z} | \tilde{\boldsymbol{r}}_\perp \rangle = \delta^2\left(\boldsymbol{r}_\perp(\theta\,(z, z_0)) - \tilde{\boldsymbol{r}}_\perp\right)\,. \tag{E3}$$

Appendix F

Green's Function for a Time-Dependent Quadratic Hamiltonian

Let it be required to compute the Green's function for a system in one dimension obeying the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = \hat{H}(t) |\Psi(t)\rangle \,, \tag{F1}$$

with a time-dependent Hamiltonian of the form

$$\hat{H}(t) = A(t)\hat{p}_x^2 + B(t)\left\{x\hat{p}_x + \hat{p}_x x\right\} + C(t)x^2 \,, \tag{F2}$$

where $A(t)$, $B(t)$ and $C(t)$ are real functions of t. Using the Magnus formula (A6)) we can write

$$|\psi(t)\rangle = \hat{U}(t, t')|\psi(t')\rangle \tag{F3}$$

with

$$\begin{aligned}
&\hat{U}(t, t') \\
&\quad = \exp\left\{ -\frac{i}{\hbar} \int_{t'}^{t} dt_1 \, \hat{H}(t_1) \right. \\
&\qquad + \frac{1}{2}\left(-\frac{i}{\hbar}\right)^2 \int_{t'}^{t} dt_2 \int_{t'}^{t_2} dt_1 \left[\hat{H}(t_2) \,,\, \hat{H}(t_1)\right] \\
&\qquad + \frac{1}{6}\left(-\frac{i}{\hbar}\right)^3 \int_{t'}^{t} dt_3 \int_{t'}^{t_3} dt_2 \int_{t'}^{t_2} dt_1
\end{aligned}$$

$$\left\{\left[\left[\hat{H}(t_3)\,,\,\hat{H}(t_2)\right],\hat{H}(t_1)\right] + \left[\left[\hat{H}(t_1)\,,\,\hat{H}(t_2)\right],\hat{H}(t_3)\right]\right\}\ldots\Big\}\,, \tag{F4}$$

and the required Green's function is given by

$$G(x,t;x',t') = \langle x|\hat{U}(t,t')|x'\rangle \tag{F5}$$

such that

$$\Psi(x,t) = \int dx\, G(x,t;x',t')\Psi(x',t')\,. \tag{F6}$$

From the fact that the operators $(x^2, \hat{p}_x^2, \{x\hat{p}_x + \hat{p}_x x\})$ are closed under commutation, leading to the Lie algebra,

$$\begin{aligned}
\left[x^2\,,\,\hat{p}_x^2\right] &= 2i\hbar\left\{x\hat{p}_x + \hat{p}_x x\right\} \\
\left[x^2\,,\,\{x\hat{p}_x + \hat{p}_x x\}\right] &= 4i\hbar x^2 \\
\left[\{x\hat{p}_x + \hat{p}_x x\}\,,\,\hat{p}_x^2\right] &= 4i\hbar\hat{p}_x^2\,,
\end{aligned} \tag{F7}$$

it is clear that $\hat{U}(t,t')$ in (F4) can be written in the form

$$\begin{aligned}
\hat{U}(t,t') = \exp\Big\{-\frac{i}{\hbar}\Big[a(t,t')\hat{p}_x^2 + \\
b(t,t')\left\{x\hat{p}_x + \hat{p}_x x\right\} + c(t,t')x^2\Big]\Big\}\,,
\end{aligned} \tag{F8}$$

where $a(t,t')$, $b(t,t')$ and $c(t,t')$ are infinite series expressions in terms of $A(t)$, $B(t)$ and $C(t)$. The precise form of (F8) can be obtained as follows. Substituting the relation

$$|\psi(t)\rangle = \hat{U}(t,t')|\psi(t')\,\rangle \tag{F9}$$

in (F1) it is seen that

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t') = \hat{H}(t)\hat{U}(t,t'), \qquad \hat{U}(t',t') = \hat{I}\,. \tag{F10}$$

This implies that

$$i\hbar\frac{\partial}{\partial t}\Big(\exp\Big\{-\frac{i}{\hbar}\Big[a(t,t')\hat{p}_x^2+$$

$$
\begin{aligned}
& b(t,t')\left\{x\hat{p}_x+\hat{p}_x x\right\}+c(t,t')x^2\Big]\Big\}\Big) \\
&= \left(A(t)\hat{p}_x^2+B(t)\left\{x\hat{p}_x+\hat{p}_x x\right\}+C(t)x^2\right) \\
&\times\left(\exp\left\{-\frac{\mathrm{i}}{\hbar}\Big[a(t,t')\hat{p}_x^2+\right.\right. \\
&\left.\left. b(t,t')\left\{x\hat{p}_x+\hat{p}_x x\right\}+c(t,t')x^2\Big]\right\}\right). \qquad \text{(F11)}
\end{aligned}
$$

The algebra in (F7) can now be used to to relate a, b and c with A, B and C. Following Wolf [40], we shall spell out these relations as follows: with the parameterization

$$
\begin{aligned}
a &= \frac{\varphi\beta}{2\sin\varphi}, \quad b=\frac{\varphi(\alpha-\delta)}{4\sin\varphi}, \quad c=-\frac{\varphi\gamma}{2\sin\varphi}, \\
&\cos\varphi=\frac{1}{2}(\alpha+\delta), \qquad \text{(F12)}
\end{aligned}
$$

it is seen that $\alpha(t,t'),\beta(t,t'),\gamma(t,t')$ and $\delta(t,t')$ satisfy the equations

$$
\begin{aligned}
& A\ddot{\alpha}-\dot{A}\dot{\alpha}+\left[4A(AC-B^2)-2A\dot{B}+2\dot{A}B\right]\alpha=0 \\
& \quad \alpha(t',t')=1, \qquad \dot{\alpha}(t',t')=2B(t') \qquad \text{(F13)} \\
& \alpha\ddot{\beta}-\beta\ddot{\alpha}-2\dot{A}=0 \\
& \quad \beta(t',t')=0, \qquad \dot{\beta}(t',t')=2A(t') \qquad \text{(F14)} \\
& \gamma=\frac{\dot{\alpha}}{2A}-\frac{\alpha B}{A} \qquad \text{(F15)} \\
& \alpha\delta-\beta\gamma=1, \qquad \text{(F16)}
\end{aligned}
$$

where $\dot{f}=\frac{df}{dt}$ and $\ddot{f}=\frac{d^2f}{dt^2}$. It is thus clear that, given $A(t)$, $B(t)$ and $C(t)$ in $\hat{H}(t)$, equations (F13)-(F16) provide $\alpha(t,t')$, $\beta(t,t')$, $\gamma(t,t')$ and $\delta(t,t')$ through (F12). Hence, we can write

$$
\begin{aligned}
& \hat{U}(t,t') \\
&= \exp\left\{-\frac{\mathrm{i}}{\hbar}\left(\frac{\varphi(t,t')}{2\sin\varphi(t,t')}\right)\Big[\beta(t,t')\hat{p}_x^2\right. \\
&\left. +\frac{1}{2}\left(\alpha(t,t')-\delta(t,t')\right)\left\{x\hat{p}_x+\hat{p}_x x\right\}-\gamma(t,t')x^2\Big]\right\}
\end{aligned}
$$

(F17)

where $\alpha, \beta, \gamma, \delta$ satisfy (F13)-(F16) and $\cos\varphi = \frac{1}{2}(\alpha + \delta)$. With A, B and C being real it is seen that $\alpha, \beta, \gamma, \delta$ and φ are real implying that $\hat{U}$ is unitary.

To find the Green's function (F5), the following observation helps: $\hat{U}(t, t')$ generates a real linear canonical transformation of the conjugate pair $(x, \hat{p}_x)$ as

$$\begin{pmatrix} \hat{U}^\dagger(t,t')x\hat{U}(t,t') \\ \hat{U}^\dagger(t,t')\hat{p}_x\hat{U}(t,t') \end{pmatrix} = \begin{pmatrix} \alpha(t,t') & \beta(t,t') \\ \gamma(t,t') & \delta(t,t') \end{pmatrix} \begin{pmatrix} x \\ \hat{p}_x \end{pmatrix}. \tag{F18}$$

In other words,

$$\begin{aligned} x\hat{U}(t,t')|\psi\rangle &= \hat{U}\left(\alpha(t,t')x + \beta(t,t')\hat{p}_x\right)|\psi\rangle \\ \hat{p}_x\hat{U}(t,t')|\psi\rangle &= \hat{U}\left(\gamma(t,t')x + \delta(t,t')\hat{p}_x\right)|\psi\rangle\,, \end{aligned} \tag{F19}$$

for any $|\psi\rangle$. Writing out (F19) explicitly in terms of matrix elements it is possible to solve for $G(x, t; x', t') = \langle x|\hat{U}(t, t')|x'\rangle$ up to a multiplicative constant phase factor (see Wolf, [40], [112], for details of the solution) : The result is

$$\begin{aligned} &G(x,t;x',t') \\ &\quad = \frac{1}{\sqrt{2\pi \mathrm{i}\hbar\beta(t,t')}} \exp\left\{\frac{\mathrm{i}}{2\hbar\beta(t,t')}\left[\alpha(t,t')x'^2\right.\right. \\ &\qquad\qquad \left.\left. -2xx' + \delta(t,t')x^2\right]\right\}, \\ &\qquad\qquad\qquad \text{if } \beta(t,t') \neq 0 \end{aligned} \tag{F20}$$

and

$$\begin{aligned} G(x,t;x',t') &= \frac{\mathrm{e}^{\left(\frac{\mathrm{i}\gamma(t,t')}{2\hbar\alpha(t,t')}\right)}}{\sqrt{\alpha(t,t')}}\delta(x' - x/\alpha(t,t'))\,, \\ &\qquad \text{if } \beta(t,t') = 0\,, \end{aligned} \tag{F21}$$

where $\delta(x' - x/\alpha)$ is Dirac delta function. In the two dimensional case, if the Hamiltonian is of the form

$$\hat{H} = A(t)\hat{p}_\perp^2 + B(t)\left\{\boldsymbol{r}_\perp \cdot \hat{\boldsymbol{p}}_\perp + \hat{\boldsymbol{p}}_\perp \cdot \boldsymbol{r}_\perp\right\} + C(t)r_\perp^2, \tag{F22}$$

then, because of the independence of the corresponding x and y motions leading to the separation of the variables x and y, the above results are extended in a straightforward manner corresponding to the replacements $x^2 \longrightarrow r_\perp^2$, $\hat{p}_x^2 \longrightarrow \hat{p}_\perp^2$, $\{x\hat{p}_x + \hat{p}_x x\} \longrightarrow \{\boldsymbol{r}_\perp \cdot \hat{\boldsymbol{p}}_\perp + \hat{\boldsymbol{p}}_\perp \cdot \boldsymbol{r}_\perp\}$.

In the case of the round magnetic lens taking (t, t') as (z, z_o), one has in $\underline{\hat{\mathbf{H}}}_{o,p}$

$$A(z) = 1/2p_0\,, \qquad B(z) = 0\,, \qquad C(z) = \frac{1}{2}p_0 F(z)\,. \tag{F23}$$

Then, taking

$$\alpha(z, z_o) = g_p\,(z, z_o)\,, \qquad \beta(z, z_o) = h_p\,(z, z_o)\,/p_0\,, \tag{F24}$$

equations (F13)-(F16) become

$$g_p''\,(z, z_o) + F(z) g_p\,(z, z_o) = 0$$

$$g_p(z_o, z_o) = 1\,, \quad g_p'(z_o, z_o) = 0 \tag{F25}$$

$$h_p''\,(z, z_o) + F(z) h_p\,(z, z_o) = 0$$

$$h_p(z_o, z_o) = 0\,, \quad h_p'(z_o, z_o) = 1 \tag{F26}$$

$$\gamma(z, z_o) = p_0 g_p'\,(z, z_o) \tag{F27}$$

$$\delta(z, z_o) = \left(1 + h_p\,(z, z_o)\, g_p'\,(z, z_o)\right)/g_p\,(z, z_o)\,, \tag{F28}$$

showing that $g_p\,(z, z_o)$ and $h_p\,(z, z_o)$ are two linearly independent solutions of the paraxial equation (2.124) with the initial conditions (2.125) as required. Since $g_p\,(z, z_o)$ and $h_p\,(z, z_o)$ are a pair of solutions of the same second order differential equation (see (F25) and (F26)) it follows from the Wronskian relation, $gh' - hg' = 1$, that $\delta(z, z_o) = h_p'\,(z, z_o)$. Substituting in (F17) $\alpha = g_p$, $\beta = h_p/p_0$, $\gamma = p_0 g_p'$, $\delta = h_p'$ and $\cos\varphi = \left(g_p + h_p'\right)/2$ we get (2.120). With $2\pi\hbar = \lambda_0 p_0$ (2.121) and (2.122) follow obviously from (F20) and (F21) since $G(\boldsymbol{r}_\perp, \boldsymbol{r}_\perp') = G(x, x')G(y, y')$ where $G(y, y')$ is obtained from $G(x, x')$ by just replacing x and x' by y and y' respectively.

Bibliography

[1] P. W. Hawkes and E. Kasper, *Principles of Electron Optics*, Vols. I and II (Academic Press, London, 1989).

[2] M. Szilagyi, *Electron and Ion Optics* (Plenum, New York, 1988).

[3] M. Conte and W. W. MacKay, *An Introduction to the Physics of Particle Accelerators* (World Scientific, Singapore, 1991).

[4] W. Glaser, Elektronen und Ionenoptik. *Handbuch der Physik* Vol. 33, Ed. S. Flügge, p. 123 (Springer, Berlin, 1956).

[5] P. W. Hawkes and E. Kasper, *Principles of Electron Optics* Vol.3: *Wave Optics* (Academic Press, London and San Diego, 1994).

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

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

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

[9] R. Jagannathan, Dirac equation and electron optics, *Dirac and Feynman - Pioneers in Quantum Mechanics*, Ed. R. Dutt and A. K. Ray (Wiley Eastern, New Delhi, 1993) 75.

[10] R. Jagannathan, *Phys. Rev. A* **42**, 6674 (1990): corrigendum, *ibid* A **44**, 7856 (1991).

[11] S. A. Khan and R. Jagannathan, Theory of relativistic electron beam transport based on the Dirac equation, *Proc. of the 3rd National Seminar on Physics and Technology of Particle Accelerators and their Applications*, (Nov. 1993, Calcutta, India) Ed. S. N. Chintalapudi (IUC-DAEF, Calcutta) 102.

[12] R. Jagannathan and S. A. Khan, Quantum theory of the optics of charged particles, *Advances in Imaging and Electron Physics*, Vol. **97**, Ed. P. W. Hawkes (Academic Press, San Diego) 257 (1996).

[13] S. A. Khan and R. Jagannathan, Quantum mechanics of charged-particle beam optics: An operator approach. Presented at the JSPS-KEK International Spring School on High Energy Ion Beams – Novel Beam Techniques and their Applications, March 1994, Japan, Preprint: IMSc/94/11 (The Institute of Mathematical Sciences, Madras, March 1994).

[14] S. A. Khan and R. Jagannathan, On the quantum mechanics of charged particle beam transport through magnetic lenses, *Phys. Rev. E* **51**, 2510 (1995).

[15] A. J. Dragt, *Lie Algebraic Methods for Ray and Wave Optics* (University of Maryland Report, in preparation).

[16] A. J. Dragt and E. Forest, *Adv. Electronics and Electron Phys.* **67**, 65 (1986).

[17] 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).

[18] G. Rangarajan, A. J. Dragt and F. Neri, *Part. Accel.* **28**, 119 (1990).

[19] R. D. Ryne and A. J. Dragt, *Part. Accel.* **35**, 129 (1991).

[20] É. Forest and K. Hirata, *A Contemporary Guide to Beam Dynamics* KEK Report 92-12 (National Laboratory for High Energy Physics, Tsukuba).

[21] É. Forest, M. Berz and J. Irwin, *Part. Accel.* **24**, 91 (1989).

[22] B. W. Montague, *Phys. Rep.* **113**, 1 (1984).

[23] M. Conte, R. Jagannathan, S. A. Khan and M. Pusterla, Beam optics of the Dirac particle with anomalous magnetic moment, Preprint: IMSc/96/03/08, INFN/AE-96/08. *Part. Accel.* **56**, 99 (1996).

[24] S. A. Khan, Transport of Dirac-particle beams through magnetic quadrupoles, Preprint: IMSc/96/33 (The Institute of Mathematical Sciences, Madras, December 1996).

[25] M. Conte, A. Penzo and M. Pusterla, *Il Nuovo Cimento A* **108**, 127 (1995).

[26] R. Jagannathan and S. A. Khan, *Selected Topics in Mathematical Physics – Professor R. Vasudevan Memorial Volume*, Eds. R. Sridhar, K. Srinivasa Rao and V. Lakshminarayanan, (Allied Publishers, Delhi, 1995) 308.

[27] R. Fedele, F. Galluccio, V. I. Man'ko, G. Miele *Phys. Lett.* **A209**, 263 (1995).

[28] A. J. Dragt and J. Finn, *J. Math. Phys.* **17**, 2215 (1976).

[29] H. Feshbach and F. Villars, *Rev. Mod. Phys.* **30**, 24 (1958).

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

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

[32] M. H. L. Pryce, *Proc. Roy. Soc. Ser. A* **195**, 62 (1949).

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

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

[35] R. S. Patton, *Path Integrals From meV to MeV*, Eds. M. Gutzwiller, A. Inomata, J. R. Klauder and L. Streit (World Scientific, Singapore, 1986) 98.

[36] J. Ximen, *Adv. Electronics and electron Phys.* **81**, 231 (1991).

[37] J. Ximen, *Advances in Imaging. Electron Physics* (Academic Press, San Diego) **97**, 359 (1996)

[38] G. S. Agarwal, and R. Simon, *Optics Comm.* **110**, 23 (1994).

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

[40] K. B. Wolf, *SIAM J. Appl. Math.* **40**, 419 (1981).

[41] W. Glaser, *Grundlagen der Elektronenoptik* (Springer, Vienna 1952).

[42] W. Glaser and P. Schiske, *Ann. Physik* **12**, 240, 267 (1953).

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

[44] H. Busch, *Arch. Elektrotech.* **18**, 583 (1927).

[45] O. Scherzer, *Z. Physik* **101**, 593 (1936).

[46] O. Scherzer, *J. Appl. Phys.* **20**, 20 (1949).

[47] M. Month and S. Turner, (Eds.) *Frontiers of Particle Beams : Observation, Diagnosis and Correction* (Springer-Verlag, Berlin, Heidelberg, 1989).

[48] H. Wiedemann, *Particle Accelerator Physics : Basic Principles and Linear Beam Dynamics* (Springer-Verlag, Berlin, Heidelberg, 1993)

[49] H. Wiedemann, *Particle Accelerator Physics II : Nonlinear and Higher-Order Beam Dynamics* (Springer-Verlag, Berlin, Heidelberg, 1995)

[50] A. Rubinowicz, *Acta Phys. Polon.* **3** 143 (1934).

[51] A. Rubinowicz, *Die Beugungswelle in der Kirchhoffschen Theorie der Beugung* (Państwowe Wydawnictwo Nauk-owe, Warsaw, 1957); 2nd ed. (PWN, Warsaw and Springer, Berlin and New York, 1966).

[52] A. Rubinowicz, *Acta Phys. Polon.* **23**, 727 (1963).

[53] A. Rubinowicz, The Miyamoto-Wolf diffraction wave. *Prog. Opt.* **4**, 199 (1965) (See Part II, Section 3: Diffraction wave in the Kirchhoff theory of Dirac electron waves).

[54] E. Durand, *C. R. Acad. Sci. Paris* **236**, 1337 (1953).

[55] Phan-Van-Loc *C. R. Acad. Sci. Paris* **237**, 649 (1953).

[56] Phan-Van-Loc, *C. R. Acad. Sci. Paris* **238**, 2494 (1954).

[57] Phan-Van-Loc, *Ann. Fac. Sci. Univ. Toulouse* **18**, 178 (1955).

[58] Phan-Van-Loc, Interprétation physique de l'expression mathématique du principe de Huygens en theérie de l'électron de Dirac. Cahiers de Physique No. 97, **12**, 327 (1958).

[59] Phan-Van-Loc, *C. R. Acad. Sci. Paris* **246**, 388 (1960).

[60] Phan-Van-Loc, *Principes de Huygens en Théorie de l'El- ectron de Dirac.* Thèse, Toulouse (1960).

[61] L. H. Thomas, *Phil. Mag.* **3**, 1 (1927).

[62] V. Bargmann, L. Michel and V. L. Telegdi, *Phys. Rev. Lett.* **2**, 435 (1959).

[63] A. A. Sokolov and I. M. Ternov, *Radiation from Relativistic Electrons*, Translated by S. Chomet and edited by C. W. Kilmister (American Institute of Physics, New York, 1986).

[64] I. M. Ternov, *Physics - Uspekhi* **38**, 409 (1995).

[65] J. S. Bell and J. M. Leinass, *Nucl. Phys. B*, **284** 488 (1987).

[66] L. N. Hand, and A. Skuja, *Proc. Conf. on Spin and Polarization Dynamics in Nuclear & Particle Physics*, Trieste, January 1988, Eds. A. O. Barut, Y. Onel and A. Penzo, (World Scientific, Singapore, 1990) 185.

[67] J. D. Jackson, *Rev. Mod. Phys.*, **48**, 417 (1976).

[68] I. M. Ternov, *Sov. Phys. JETP*, **71**, 654 (1990).

[69] H. C. Corben, *Classical and Quantum Theories of Spinning Particle* (Holden-Day, Inc., San Francisco, 1968).

[70] Ya. S. Derbenev and A. M. Kondratenko, *Sov. Phys. JETP*, **37**, 968 (1973).

[71] K. Heinemann, On Stern-Gerlach forces allowed by special relativity and the special case of the classical spinning particle of Derbenev-Kondratenko, *e-print:* physics/9611001

[72] D. P. Barber, K. Heinemann and G. Ripken, *Z. Phys. C*, **64**, 117 (1994).

[73] D. P. Barber, K. Heinemann and G. Ripken, *Z. Phys. C*, **64**, 143 (1994).

[74] Y. Onel, A. Penzo and R. Rossmanith, *AIP Conf. Proc.* 150 Eds. R. G. Lernerand and D. F. Geesaman, (American Institute of Physics, New York, 1986) 1229.

[75] T. Niinikoski and R. Rossmanith, *Nucl. Instrum. Methods A*, **255**, 46 (1987).

[76] M. Conte, A. Penzo, A. Pisent and M. Pusterla, Analytical treatment of the spin splitter, Internal Report : INFN/TC-88/25.

[77] M. Conte, *Proc. Conf. on Spin and Polarization Dynamics in Nuclear & Particle Physics*, Trieste, January 1988, Eds. A. O. Barut, Y. Onel and A. Penzo (World Scientific, Singapore, 1990) 316.

[78] A. Pisent, *Proc. Conf. on Spin and Polarization Dynamics in Nuclear & Particle Physics*, Trieste, January 1988, Eds. A. O. Barut, Y. Onel and A. Penzo (World Scientific, Singapore, 1990) 327.

[79] M. Conte, and M. Pusterla, *Il Nuovo Cimento A*, **103**, 1087 (1990).

[80] M. Conte, Y. Onel, A. Penzo, A. Pisent, M. Pusterla and R. Rossmanith, The spin-splitter concept, Internal Report : INFN/TC-93/04.

[81] J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, New York, 1962).

[82] K. Hagedorn, *Relativistic Kinematics* (W. A. Benjamin, New York, 1963).

[83] V. M. Castaño, *Optik* **81**, 35 (1988).

[84] V. M. Castaño, *Computer Simulation of Electron Microscope Diffraction and Images*, Eds. W. Krakow and M. A. O'Keefe (Minerals, Metals and Materials Society, Warrendale, PA, 1989) 33.

[85] V. M. Castaño, G. Vázquez-Polo and R. Gutierrez-Castrejón, *Signal and Image Processing in Microscopy and Microanalysis. Proc. of the 10th Pfefferkorn Conference*, Cambridge, 1991, Eds. P. W. Hawkes, W. O. Saxton and M. A. O'Keefe (Scanning Microsc. Suppl. 6 (published 1994)) 414.

[86] G. V. Polo, D. R. Acosta and V. M. Castaño, *Optik* **89**, 181 (1992).

[87] E. P. Wigner, *Phys. Rev. Lett.* **40**, 749 (1932).

[88] M. Hillery, R. F. O'Connell, M. O. Scully and E. P. Wigner, *Phys. Rep.* **106** 121 (1984).

[89] N. Z. Balazs and B. K. Jennings, *Phys. Rep.* **104**, 347 (1984).

[90] V. I. Tatarskiĭ, *Sov. Phys. Usp.* **26**, 311 (1983).

[91] P. Carruthers and F. Zachariassen, *Rev. Mod. Phys.* **55** 245 (1983).

[92] N. Mukunda, *Pramana* **11**, 1 (1978).

[93] K. H. Mariwalla, *Proc. Conference on Probability, Stochastic Process and Statistical Mechanics*, MATSCIENCE Report No.97 (The Institute of Mathematical Sciences, Madras, 1979)

[94] R. Jagannathan, R. Simon, E. C. G. Sudarshan and R. Vasudevan, *Phys. Lett.* **A120**, 161 (1987).

[95] M. J. Bastians, *Optik* **82**, 173 (1989).

[96] R. Simon, E. C. G. Sudarshan and N. Mukunda, *Phys. Rev.* **A36**, 3868 (1987).

[97] F. Lenz in *Magnetic Electron Lenses*, Topics in Current Physics Vol. **18** ed. P. W. Hawkes, (Springer-Verlag, Berlin 1982).

[98] B. R. Mallow, *Phys. Rev.* **162**, 1256 (1967).

[99] A. Grossmann, *Comm. Math. Phys.* **48**, 191 (1976).

[100] A. Royer, *Phys. Rev.* **15**, 449 (1977).

[101] G. Dattoli, C. Mari and A. Torre, *A note on Wigner phase space and beam transport*, Report No. RI/INN/91/30 ENEA, Rome.

[102] G. R. Shin, I. Bialynicki-Birula and J. Rafelski, *Phys. Rev.* **A46**, 645 (1982).

[103] H. T. Elze and U. Heinz, *Phys. Rep.* **183**, 81 (1989).

[104] J. Liñares, Optical propagators in vector and spinor theories by path integral formalism, in *Lectures on Path Integration: Trieste 1991* ed. H. A. Cerdeira *et al.*, (World Scientific, Singapore, 1993).

[105] E. Bauer, *Rep. Prog. Phys.* **57**, 895 (1994).

[106] T. D. Newton and E. P. Wigner, *Rev. Mod. Phys.* **21**, 400 (1949).

[107] A. O. Barut, and R. Raczka, *Theory of Group Representations and Applications* (World Scientific, Singapore 1986).

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

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

[110] R. M. Wilcox, *J. Math. Phys.* **4**, 962 (1967).

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

[112] K. B. Wolf, *Integral Transforms in Science and Engineering* (Plenum, New York, 1979).