Paraxial light and atom optics: The optical Schrödinger equation and beyond

Monika A. M. Marte
Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria

Stig Stenholm
Helsinki Institute of Physics, University of Helsinki, PO Box 9, FIN-00014, Helsinki, Finland
(Received 27 February 1997)

Paraxial light and atom optics are compared. To lowest order the slowly varying amplitude of a light field in a dielectric medium with a spatially dependent refractive index satisfies an equation which has the form of a Schrödinger equation: the ''optical Schrödinger equation.'' The unsystematic procedure of neglecting certain second-order derivatives is replaced by a systematic expansion which allows the calculation of consecutive corrections. The general theory is applied to harmonic motion in a graded index fiber and to tunneling between coupled fibers. The physical relations between wave evolution of massive particles and paraxially propagating waves are elucidated. [S1050-2947(97)00710-5]

PACS number(s): 03.75.−b, 42.50.−p, 42.25.Bs, 03.65.−w

I. INTRODUCTION

Much interest in physics has recently been directed towards time-dependent phenomena connected with wave propagation. Waves appear in many forms, but the two most fundamental ones are derived from the propagation of electromagnetic radiation and the matter-wave picture of massive particles. The former are described by Maxwell's equations and the latter by the nonrelativistic Schrödinger equation. Experimentally the demand for fast optical communication has inspired investigations into pulsed propagation, and the technology of pulsed lasers has enabled the researchers to follow microscopic time evolution in atoms, molecules, and semiconductor structures.

The physical properties of matter waves differ from those of electromagnetic radiation. Schrödinger's equation leads to dispersion of the wave packets and distortion of the initial pulse shape. In Maxwell's case an initial pulse can propagate in a passive dielectric without distortion. The massless electromagnetic waves have a constant velocity independent of energy in contrast to the matter waves. In addition, the electromagnetic field is a vector quantity. It is, however, known that in the paraxial approximation [1,2], propagation along the principal axis of the wave vector can simulate the time evolution in the Schrödinger case with the position variables representing the transverse coordinates of the paraxial rays. If we, in addition, assume that the dielectric properties of the medium depend on position, the lowest-order paraxial equations are found to be identical with the wave equation of quantum mechanics; this analogy can be developed to a considerable extent [3,4].

In a homogeneous dielectric, the electromagnetic wave propagation can always be discussed from a frame moving with the velocity of light in the medium. In an inhomogeneous medium, moving along the axis of the system does not correspond exactly to the progress of time because the medium introduces a velocity dependence into the equations. Introducing a scalar paraxial wave equation destroys some of the vector properties of the radiation fields. The formulation, however, allows the interpretation of the propagation as the evolution of a massive particle, and Schrödinger-like behavior emerges. Marcuse argues that this is approximately correct, provided that the variation of the refractive index is small over the distance of the optical wavelength [5], but no estimate of the error is given. To investigate the consistency of the paraxial approximation, one must find the wave corrections to the particlelike behavior. This has been done by Lax, Louisell, and McKnight [6] (cf. also the Appendix in [7]). They find that the approach is consistent, since deviations from the simple scalar paraxial wave equation are of higher order in a small parameter; their work can form the basis for a more detailed discussion of the ''radiation corrections'' to the paraxial equations.

In this work, we compare the Schrödinger wave mechanics with the classical electromagnetic theory. From a fundamental point of view this is the correct procedure. The theories are at the same conceptual level, the wave function of a photon is just the classical electromagnetic field; this fact has recently been emphasized by Bialynicki-Birula [8]. Both wave theories can, of course, be subjected to second quantization. This has been discussed within the paraxial approximation by Deutsch and Garrison [9].

We are going to follow the work in Ref. [6] and develop a consistent approximation scheme giving the paraxial matter-wave equation in the lowest order. Because this describes the wave dynamics of a massive particle, we can discuss many of the phenomena recently investigated in atom optics, such as dispersion, interference, and tunneling. In the lowest paraxial approximation, these will appear exactly as in the Schrödinger case, except that the propagation follows the axial direction instead of time. Experimentally all of these phenomena can be investigated in optical fibers [2,10], and because of their interest to communication technology, the experiments may be both interesting and feasible. As long as the fiber can be regarded as a linear medium, it may be used for analogous simulations of various matter-wave phenomena of recent interest in atom optics [11–14].

So far the analogy is precise but also a little trivial. Interesting questions arise if we ask when and how the electro-

magnetic wave propagation starts to differ from Schrödinger propagation. This can, of course, be investigated by solving the exact Maxwell equations and comparing the result with that of the paraxial approximation. Here we take a different approach; we compute the lowest-order correction to the paraxial result. In this manner we can see directly where the approximation has to be corrected and what form the corrections will take. Because we are especially interested in these corrections, it is advantageous to evaluate them directly. By solving the corresponding Maxwell equations, we can then estimate when the approximation used becomes inadequate, which defines the range of validity of our conclusions. Nonparaxial corrections to light wave propagation are currently an active field of research (cf., for example, Ref. [15] for optical applications or Ref. [16] for a mathematical approach), but the aim of the present paper differs from these investigations in its intention to compare the propagation of light and matter waves.

In Sec. II the paraxial wave equations are derived starting from Maxwell's equations: First the paraxial approximation is carried out in the usual way, then a systematic expansion which enables the calculation of higher-order corrections is formulated. The general theory is applied to the examples of harmonic motion and tunneling in a double well in Sec. III. Section IV provides a summary and discussion of the results and an outlook to possible extensions of the theory. The full vector equations and alternative forms of the expansion are given in an Appendix.

II. SETTING UP THE PROBLEM

A. Aims and scope

We are going to consider the propagation of electromagnetic waves in a *monochromatic and paraxial approximation*; this assumes that the amplitude of the fields varies spatially much more slowly than the wave propagation in the principle direction $\mathbf{k}_0 = k_0 \mathbf{e}_z$, so that all propagation vectors transverse to this are assumed small. Thus we introduce the approximation

$$\left| \frac{\partial^2 \mathbf{E}}{\partial z^2} \right| \ll k_0 \frac{\partial \mathbf{E}}{\partial z}, \tag{1}$$

where $\mathbf{E}$ is the slowly varying electric field amplitude.

Within the approximation defined, the steady state paraxial propagation occurs in the direction

$$z \equiv \frac{\tau}{\hbar k_0 n_0}, \tag{2}$$

where n_0 is the constant bulk refractive index of the background medium. In the steady state, we can interpret the variable τ as a time and write the propagation equation in the form

$$i\hbar \frac{\partial}{\partial \tau} \psi(x,y;\tau) = -\frac{\hbar^2}{2m_{\mathrm{opt}}} \nabla_\perp^2 \psi(x,y;\tau) + V_{\mathrm{opt}}(x,y;\tau)\psi(x,y;\tau), \tag{3}$$

where $\psi(x,y;\tau)$ is the scalar amplitude of the electromagnetic field, and $\nabla_\perp^2$ is the two-dimensional transverse Laplacian; see below for a systematic derivation. It is easily seen that this equation is identical with the nonrelativistic Schrödinger equation describing the propagation of a massive particle in two dimensions, cf. Ref. [3]. The optical effective mass is found to be

$$m_{\mathrm{opt}} = (\hbar k_0 n_0)^2.$$

Thus the heavy mass limit, corresponding to semiclassical behavior, appears when the wavelength of the propagating field becomes short compared with the rate of change of the optical potential

$$V_{\mathrm{opt}} = -\frac{1}{2}\left[\frac{n^2(\mathbf{r}) - n_0^2}{n_0^2} \right]. \tag{4}$$

For a detailed discussion of the emergence of this limit, see Sec. V. The validity of the paraxial approximation does, however, impose a limitation on the optical potential. This must *necessarily* be a small perturbation on the free propagation for the condition Eq. (1) to make sense. The amplitude can be a slowly varying function only if $[n^2(x,y,\tau) - n_0^2]$ is small in a definite sense that will be specified below.

B. Maxwell theory in an inhomogeneous medium

We start our discussion from the Maxwell equations in the ordinary form

$$\frac{\partial}{\partial t}\mathbf{D}(\mathbf{r},t) = \nabla \times \mathbf{H}(\mathbf{r},t), \quad \frac{\partial}{\partial t}\mathbf{B}(\mathbf{r},t) = -\nabla \times \mathbf{E}(\mathbf{r},t),$$

$$\nabla \cdot \mathbf{D}(\mathbf{r},t) = 0, \quad \nabla \cdot \mathbf{B}(\mathbf{r},t) = 0. \tag{5}$$

The equations are supplemented by the relations

$$\mathbf{D}(\mathbf{r},t) = \varepsilon(\mathbf{r})\varepsilon_0 \mathbf{E}(\mathbf{r},t),$$

$$\mathbf{B}(\mathbf{r},t) = \mu_0 \mathbf{H}(\mathbf{r},t), \tag{6}$$

where $\varepsilon(\mathbf{r})$ gives the spatial dependence of the inhomogeneous medium (at the optical frequency ω); the magnetic polarizability of the medium is neglected. The spatially varying dielectric constant is used to define a position-dependent refractive index

$$n^2(\mathbf{r}) = \varepsilon(\mathbf{r}). \tag{7}$$

In *free space* radiation problems, we introduce the potentials $\mathbf{A}$ and Φ through the definitions

$$\mathbf{E}(\mathbf{r},t) = -\left(\frac{\partial \mathbf{A}(\mathbf{r},t)}{\partial t} + \nabla \Phi \right),$$

$$\mathbf{B}(\mathbf{r},t) = \nabla \times \mathbf{A}(\mathbf{r},t). \tag{8}$$

With the Lorentz gauge condition, the wave equations for the two potentials separate, and in the Coulomb gauge all radiation effects can be restricted to the transverse components of

the electromagnetic fields. Then the vector potential **A** contains all radiation effects and the scalar potential Φ describes the electrostatic interactions.

In an inhomogeneous medium the choice

$$\nabla\cdot\mathbf{A}=0,$$

$$\Phi=0 \tag{9}$$

is not consistent, even when there are no external charges present, except in some very special cases, some of which are discussed below. However, for monochromatic radiation propagating through an inhomogeneous dielectric it is still possible to eliminate one of the electromagnetic field vectors, as in the vacuum case, but the ensuing equations contain additional terms deriving from the space variation of the medium. We find

$$\left(\nabla^2+\varepsilon(\mathbf{r})\frac{\omega^2}{c^2}\right)\mathbf{E}=\nabla(\nabla\cdot\mathbf{E})=-\nabla[\nabla(\ln\varepsilon(\mathbf{r}))\cdot\mathbf{E}] \tag{10}$$

and

$$\left(\nabla^2+\varepsilon(\mathbf{r})\frac{\omega^2}{c^2}\right)\mathbf{H}=-\nabla(\ln\varepsilon(\mathbf{r}))\times[\nabla\times\mathbf{H}(\mathbf{r},t)]. \tag{11}$$

For the second form of Eq. (10) we have made use of the third of Maxwell's equations (5),

$$\nabla\cdot\mathbf{E}=-\nabla(\ln\varepsilon(\mathbf{r}))\cdot\mathbf{E}. \tag{12}$$

Since a nonzero divergence of **E** leads to a coupling between the components of the electric field vector, these equations are much harder to solve than the corresponding vacuum equations. But certain simple cases which are characterized by

$$\mathbf{E}\perp\nabla\varepsilon(\mathbf{r}), \tag{13}$$

resulting in a *scalar* wave equation for the electric field [17], are easily treated. We proceed to discuss these special cases and relegate the general vector case to the first part of the Appendix.

C. Stratified media

Many physical systems, such as the atmosphere and technologically fabricated dielectric devices, possess a layered structure [18], such as, for example, planar waveguides; that is, the dielectric properties vary only in one direction, as depicted in Fig. 1. For $\varepsilon=\varepsilon(x)$, for instance, the gradient of the dielectric constant becomes

$$\nabla\varepsilon(\mathbf{r})=\varepsilon'\mathbf{e}_x. \tag{14}$$

This allows several simple solutions as discussed in Refs. [2,10,19,20].

1. E-wave solution

First we investigate a solution with the electric field vector linearly polarized in a direction transverse to the gradient of the dielectric properties and set

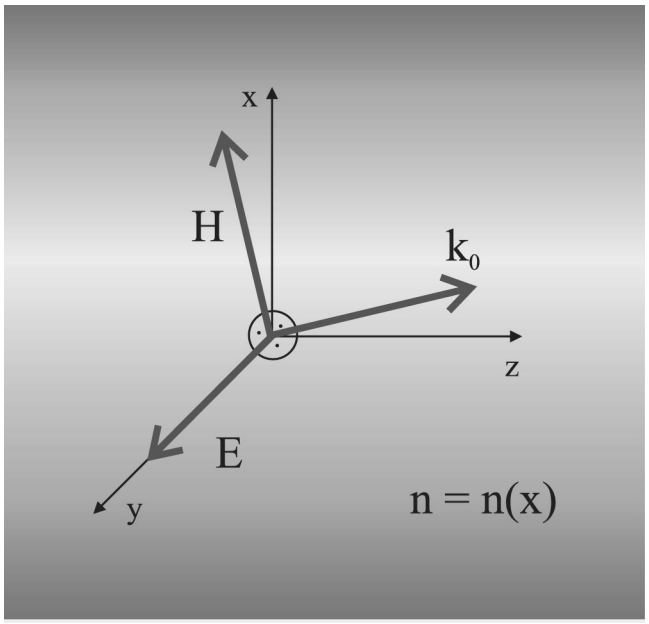

FIG. 1. Polarization configuration that gives rise to a scalar wave equation for the E field in a plane-stratified medium.

$$\mathbf{E}(x,z)=\mathbf{e}_y\mathcal{E}(x)e^{i\kappa z}. \tag{15}$$

Here, as throughout, z is the propagation direction. Obviously this field is transverse, and the field amplitude $\mathcal{E}(x)$ is from Eq. (10) found to satisfy

$$\left[\frac{d^2}{dx^2}+\left(\varepsilon(x)\frac{\omega^2}{c^2}-\kappa^2\right)\right]\mathcal{E}(x)=0, \tag{16}$$

which is clearly of the form of a time-independent Schrödinger equation. With this solution, the magnetic field vector cannot be chosen transverse, but becomes

$$\mathbf{H}(x,z)=-\frac{1}{\omega\mu_0}[i\mathcal{E}'(x)\mathbf{e}_z+\kappa\mathcal{E}(x)\mathbf{e}_x]e^{i\kappa z}. \tag{17}$$

This solution is sometimes also referred to as a TE mode [2,10]. It allows one to describe the propagation of the radiation in terms of transverse fields, and it leads to simple Schrödinger-like behavior of the field amplitudes. In our approximate expansion later in the paper we will concentrate on this type of solution. In order to see the limitations of this approach, we briefly consider some other situations too.

2. H-wave solution

In this case we choose the magnetic field transverse to the gradient of the dielectric,

$$\mathbf{H}(x,z)=\mathbf{e}_y\mathcal{H}(x)e^{i\kappa z}. \tag{18}$$

Equation (11) now becomes

$$\frac{d}{dx}\left(\frac{1}{\varepsilon(x)}\frac{d\mathcal{H}(x)}{dx}\right)+\left(\frac{\omega^2}{c^2}-\frac{\kappa^2}{\varepsilon(x)}\right)\mathcal{H}(x)=0. \tag{19}$$

This is not of the form of a simple Schrödinger equation, but is a similar eigenvalue problem. The electric field vector is now found to be

$$\mathbf{E}(x,z)=\frac{1}{\varepsilon_0\varepsilon(x)\omega}[i\mathcal{H}'(x)\mathbf{e}_z+\kappa\mathcal{H}(x)\mathbf{e}_x]e^{i\kappa z}. \quad (20)$$

This solution is also called a TM mode.

From Eq. (20) we can see that it is not always possible to choose the electric field vector such that it is transverse to the direction of propagation. Because of the x dependence of $\varepsilon(x)$, the divergence of $\mathbf{E}$ is found to satisfy Eq. (12). For a wave propagating with a longitudinal field component, no simple gauge allows a separation of the evolution equations of the two potential functions.

3. Normal incidence

A trivial solution in the stratified medium can be found when the propagation coincides with the direction of the gradient of the dielectric. Then we can choose the solution in the form

$$\mathbf{E}(z)=\mathbf{e}_y\mathcal{E}_0\exp[i\varphi(z)], \quad (21)$$

where $\varphi(z)$ satisfies an equation of the Riccatti type,

$$i\varphi''(z)-\varphi'(z)^2+\frac{\omega^2}{c^2}\varepsilon(z)=0. \quad (22)$$

Assuming slow spatial variation and neglecting the second derivative, we obtain a solution of the eikonal type. The calculations in the Appendix also utilize the freedom to choose the function $\varphi(\mathbf{r})$ in an arbitrary way to eliminate some part of the spatial dependence of the solution.

D. Cylindrical symmetry

In cylindrical waveguides and optical fibers, the refractive index often changes radially only. If we choose the axis of the cylinder to be the propagation direction z, we have $\varepsilon=\varepsilon(\sqrt{x^2+y^2})$. Such a symmetrical waveguide provides, besides the TE waves in planar waveguides, another (two-dimensional) example allowing special solutions which satisfy a scalar wave equation. As before, we seek a transverse solution with $\mathcal{E}_z=0$, i.e.,

$$\mathbf{E}(x,y,z)=[\mathcal{E}_x(x,y)\mathbf{e}_x+\mathcal{E}_y(x,y)\mathbf{e}_y]e^{i\kappa z}. \quad (23)$$

Fields with zero component along the radial unit vector $\mathbf{e}_r=\cos\phi\,\mathbf{e}_x+\sin\phi\,\mathbf{e}_y$ in cylindrical coordinates (r,ϕ,z),

$$\mathcal{E}_r=\mathcal{E}_x(x,y)\cos\phi+\mathcal{E}_y(x,y)\sin\phi=0, \quad (24)$$

are orthogonal to the gradient of $\varepsilon(r)$; from Eq. (12) it follows immediately that the components $\mathcal{E}_x(x,y)$ and $\mathcal{E}_y(x,y)$ then decouple, each satisfying the scalar Eq. (16). We thus choose

$$\mathbf{E}(r,z)=\mathcal{E}_\phi(r)\mathbf{e}_\phi e^{i\kappa z}, \quad (25)$$

where $\mathbf{e}_\phi$ is the unit tangential vector $\mathbf{e}_\phi=-\sin\phi\,\mathbf{e}_x+\cos\phi\,\mathbf{e}_y$. The two scalar equations for the transverse Cartesian coordinates can of course be combined into a single equation for the tangential component $\mathcal{E}_\phi(r)$,

$$\left[\frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right)+\left(\varepsilon(r)\frac{\omega^2}{c^2}-\kappa^2\right)-\frac{1}{r^2}\right]\mathcal{E}_\phi(r)=0. \quad (26)$$

The magnetic field has radial and longitudinal components,

$$\mathbf{H}(r,z)=-\frac{1}{\omega\mu_0}\left[i\frac{1}{r}\frac{d}{dr}[r\mathcal{E}_\phi(r)]\mathbf{e}_z+\kappa\mathcal{E}_\phi(r)\mathbf{e}_r\right]e^{i\kappa z}. \quad (27)$$

For weakly confining cylindrical wave guides with a very weak rate of change of the refractive index, it is even possible to approximate the solution by a linearly polarized TE wave [10].

III. A SYSTEMATIC EXPANSION

A. The perturbation equations

From now on we will concentrate on wave packets consisting of superpositions of TE-type solutions propagating, e.g., in a stratified medium or in a circularly symmetric fiber with an appropriately chosen transverse refractive index profile $n(\mathbf{r}_\perp)$. Let us denote the generic component of the electric field satisfying the scalar wave equation by $E(\mathbf{r})$. Pulling out a ''carrier plane wave'' propagating in the z direction we write

$$E(\mathbf{r})=\exp[ik_0n_0z-i\omega t]\psi(\mathbf{r}), \quad (28)$$

where the mean propagation vector in vacuum is introduced as

$$k_0=\frac{\omega}{c} \quad (29)$$

and n_0 stands for the bulk refractive index of the medium. From Eq. (10) the slowly varying amplitude $\psi(\mathbf{r})$ is found to satisfy the equation

$$(\nabla+ik_0n_0\mathbf{e}_z)^2\psi(\mathbf{r})-\frac{n^2(\mathbf{r}_\perp)\omega^2}{c^2}\psi(\mathbf{r})=0. \quad (30)$$

This equation is the starting point for our approximations.

We next introduce a small parameter, the angle Θ between the $\mathbf{k}$ vector and the vector $\mathbf{k}_0=k_0\mathbf{e}_z$. Defining

$$\Theta=\angle(\mathbf{k}_0,\mathbf{k}),\quad \mathbf{k}=k_0+\mathbf{q} \quad (31)$$

we get

$$\Theta\sim|\sin\Theta|=\frac{|\mathbf{q}_\perp|}{|\mathbf{k}_0+\mathbf{q}|},\quad \cos\Theta=\frac{|k_0+q_z|}{|\mathbf{k}_0+\mathbf{q}|}. \quad (32)$$

Thus the transverse deviations of the wave vector scale as Θ, and the longitudinal ones as Θ^2.

In order to obtain a series expansion in Θ, we introduce the appropriately scaled variables:

$$\tilde{\mathbf{r}}_\perp=\Theta k_0n_0\mathbf{r}_\perp, \quad (33)$$

$$\tilde{z}=\Theta^2 k_0 n_0 z\equiv\Theta^2\tau/\hbar. \tag{34}$$

In the latter expression for $\tilde{z}$ we have used the definition (2). Thus the plane wave exponential scales as

$$\exp[ik_0 n_0 z]=\exp[i\tilde{z}|\Theta^2], \tag{35}$$

which gives rise to fast oscillations over ranges of $\tilde{z}$ of the order unity. Introducing these scaled variables into Eq. (30) we obtain

$$n^2(\mathbf{r}_\perp)k_0^2\psi-n_0^2k_0^2\psi+2i\Theta^2n_0^2k_0^2\frac{\partial\psi}{\partial\tilde{z}}+\Theta^2n_0^2k_0^2\tilde{\nabla}_\perp^2\psi+\Theta^4\frac{\partial^2\psi}{\partial\tilde{z}^2}=0. \tag{36}$$

Rearranging the terms, we obtain the equation in a form suitable for a perturbation treatment,

$$i\frac{\partial\psi}{\partial\tilde{z}}=\left\{-\frac{1}{2}\tilde{\nabla}_\perp^2-\frac{1}{2\Theta^2}\left[\frac{n^2(\tilde{\mathbf{r}}_\perp)-n_0^2}{n_0^2}\right]\right\}\psi-\frac{\Theta^2}{2}\frac{\partial^2\psi}{\partial\tilde{z}^2}. \tag{37}$$

When the term of order $O(\Theta^2)$ is neglected, we find the Schrödinger-like equation of the paraxial approximation. In order for this to be consistent, we require that the deviations of the refractive index from its bulk value n_0 must be small in such a manner that

$$\left[\frac{n^2(\tilde{\mathbf{r}}_\perp)-n_0^2}{n_0^2}\right]=O(\Theta^2). \tag{38}$$

This assumption is similar to the one made in Ref. [6].

When we neglect the last term in Eq. (37) and restore the original space variables and the ''time'' τ, we find the optical Schrödinger equation (3) with the ''potential'' given as in Eq. (4). Because of the scaling (38) we introduce the effective potential

$$\tilde{V}_{\rm opt}=-\frac{1}{2\Theta^2}\left[\frac{n^2(\tilde{\mathbf{r}}_\perp)-n_0^2}{n_0^2}\right], \tag{39}$$

which is assumed to be of order $\Theta^0\sim 1$.

B. The iterated solution

In order to determine the lowest-order corrections explicitly, we expand the ''wave function'' $\psi(\tilde{\mathbf{r}}_\perp;\tilde{z})$ in a power series of Θ,

$$\psi(\tilde{\mathbf{r}}_\perp;\tilde{z})=\sum_{\mu=0}^{\infty}\Theta^{2\mu}\ \psi^{(2\mu)}(\tilde{\mathbf{r}}_\perp;\tilde{z}). \tag{40}$$

From Eq. (36) it is obvious that only even powers of Θ are required here; this is a special simplification holding only for solutions of the scalar type, where the right-hand side of Eq. (10), which introduces odd orders, is zero (cf. the section on the general vector case in the Appendix). Insertion and comparison of the coefficients $\propto\Theta^{2\mu}$ directly yields

TABLE I. The first few ''optical corrections'' as a function of $\alpha(\tilde{z})\equiv-(i/2)\tilde{z}\tilde{E}_\nu^2$.

Order	Additive correction term $\mathcal{C}^{(2\mu)}(\tilde{z})\equiv c_\nu^{(2)}(\tilde{z})/c_\nu^{(0)}$
$\mu=1$	$\mathcal{C}^{(2)}(\tilde{z})=\alpha(\tilde{z})$
$\mu=2$	$\mathcal{C}^{(4)}(\tilde{z})=\alpha(\tilde{z})\tilde{E}_\nu+\frac{\alpha(\tilde{z})^2}{2!}$
$\mu=3$	$\mathcal{C}^{(6)}(\tilde{z})=\alpha(\tilde{z})\frac{5}{4}\tilde{E}_\nu^2+\frac{\alpha(\tilde{z})^2}{2!}2\tilde{E}_\nu+\frac{\alpha(\tilde{z})^3}{3!}$
$\mu=4$	$\mathcal{C}^{(8)}(\tilde{z})=\alpha(\tilde{z})\frac{7}{4}\tilde{E}_\nu^3+\frac{\alpha(\tilde{z})^2}{2!}\frac{7}{2}\tilde{E}_\nu^2+\frac{\alpha(\tilde{z})^3}{3!}3\tilde{E}_\nu+\frac{\alpha(\tilde{z})^4}{4!}$

$$\left[i\frac{\partial}{\partial\tilde{z}}-H_\perp\right]\psi^{(0)}(\tilde{\mathbf{r}}_\perp;\tilde{z})=0,\quad \mu=0 \tag{41}$$

$$\left[i\frac{\partial}{\partial\tilde{z}}-H_\perp\right]\psi^{(2\mu)}(\tilde{\mathbf{r}}_\perp;\tilde{z})=-\frac{1}{2}\frac{\partial^2}{\partial\tilde{z}^2}\psi^{(2\mu-2)}(\tilde{\mathbf{r}}_\perp;\tilde{z}),\qquad \mu=1,2,3,\ldots \tag{42}$$

with

$$H_\perp=-\frac{1}{2}\tilde{\nabla}_\perp^2+\tilde{V}_{\rm opt} \tag{43}$$

and $\tilde{V}_{\rm opt}$ defined in Eq. (39). We see that only the lowest-order amplitude $\psi^{(0)}$ is a solution of a proper Schrödinger equation (41); the higher-order ''optical corrections'' are determined by the inhomogeneous equations (42).

Now we solve the above hierarchy of equations by determining the eigenbasis $\{\tilde{\varphi}_\nu^S(\tilde{\mathbf{r}}_\perp)\}$ of $H_\perp$ and expanding the zero-order amplitude $\psi^{(0)}(\tilde{\mathbf{r}}_\perp;\tilde{z})$ in this basis,

$$\psi^{(0)}(\tilde{\mathbf{r}}_\perp;\tilde{z})=\sum_\nu\ c_\nu^{(0)}e^{-i\tilde{E}_\nu\tilde{z}}\tilde{\varphi}_\nu^S(\tilde{\mathbf{r}}_\perp), \tag{44}$$

with ''time''-independent coefficients $c_\nu^{(0)}$. Since the Hamiltonian $H_\perp$ is ''time'' independent in the present case, it commutes with $\partial^2/\partial\tilde{z}^2$, which means that the right-hand side of Eq. (42) also represents a solution of the homogeneous equation (41). Setting

$$\psi^{(2\mu)}(\tilde{\mathbf{r}}_\perp;\tilde{z})=\sum_\nu\ c_\nu^{(2\mu)}(\tilde{z})e^{-i\tilde{E}_\nu\tilde{z}}\tilde{\varphi}_\nu^S(\tilde{\mathbf{r}}_\perp),\quad \mu\neq 0 \tag{45}$$

with $\tilde{z}$-dependent coefficients, we get the recurrence relation

$$\dot{c}_\nu^{(2\mu)}(\tilde{z})=\frac{i}{2}\ddot{c}_\nu^{(2\mu-2)}(\tilde{z})+\tilde{E}_\nu\dot{c}_\nu^{(2\mu-2)}(\tilde{z})-\frac{i}{2}\tilde{E}_\nu^2c_\nu^{(2\mu-2)}(\tilde{z}), \tag{46}$$

where dots represent differentiation with respect to $\tilde{z}$. This can easily be solved by iteration, assuming the ''initial condition'' $c_\nu^{(2\mu)}(\tilde{z}=0)=0$ for $\mu>0$; the first few orders are given in Table I.

Returning to unscaled variables and defining

$$\tilde{E}_\nu = E_\nu^S/\Theta^2, \tag{47}$$

which gives rise to the notational identities

$$\tilde{E}_\nu \tilde{z} \equiv E_\nu^S k_0 n_0 z \equiv E_\nu^S \tau/\hbar, \tag{48}$$

we get, to first order,

$$\psi(\mathbf{r}_\perp;\tau) = \sum_\nu c_\nu^{(0)} e^{-iE_\nu^S \tau/\hbar} \varphi_\nu(\mathbf{r}_\perp) + \sum_\nu c_\nu^{(0)} \left(-\frac{i}{2}\right) E_\nu^{S^2} \frac{\tau}{\hbar} e^{-iE_\nu^S \tau/\hbar} \varphi_\nu(\mathbf{r}_\perp) + \cdots. \tag{49}$$

Note that, due to the definition Eq. (2), $\tau/\hbar$ and therefore also E_ν^S are dimensionless. The second term immediately gives a rough estimate for the typical propagation ''times'' τ after which the validity of the truncation after the lowest-order correction term breaks down; this happens for times of the order

$$\tau_{\mathrm{corr}}/\hbar = 2/(E_{\nu_0}^S)^2, \tag{50}$$

with $E_{\nu_0}^S$ being a typical eigenvalue of the problem. For $\tau \approx \tau_{\mathrm{corr}}$ the modulus of the first-order correction coefficient approaches unity; clearly for such propagation times truncation after the first-order correction fails and more terms have to be kept. Numerically it has turned out to be useful to take the eigenvalue belonging to the eigenfunction of maximum overlap with the initial wave function as a typical choice for $E_{\nu_0}^S$. From Table I we infer that the correction coefficients grow according to power laws in τ, $C^{(2\mu)} \propto \tau^\mu$. Thus the truncation of the infinite series makes sense only provided that the modulus of the highest-order coefficient taken into account is considerably less than unity; see the discussion about the validity of the numerical results at the end of the next section.

Equations (41) and (42) give an example of a set of singular perturbation equations; the zeroth-order solution turns up as a singularity in the inversion required in all subsequent orders. Such perturbations are known from many problems in physics, starting with the secular perturbation theory of planetary motion. A straightforward computation of the correction terms leads to unbounded terms in the time evolution, just as in the present case. These limit the range of validity of the result and fail to describe the long time behavior correctly. Applied mathematics has developed a variety of methods to treat such systems; the most efficient one is based on a multiple-time-scale expansion [21].

The present problem shows many similarities with the mathematical model for the operation of the free-electron laser [22]. Here the appearance of nonsecular unbounded terms does not cause any complications, because the gain is evaluated for a limited time period only. It has, however, been shown [23] that the use of a multiple-time-scale expansion allows one to construct a perturbation result that gives a good approximation for both short and long time periods.

In this work we are satisfied with applying the straightforward perturbation approach leading to unbounded secular terms. The reason is that we want to evaluate the validity of the zeroth-order solution and compare this to the exact electromagnetic propagation following from Maxwell's equations. It is then expedient to have an expression for the first-order correction, which can be evaluated and compared both with the zeroth-order solution and the correct result. It seems pointless to attempt a valid long time expansion, when the exact result can be obtained from the computation of Maxwell's equations.

IV. APPLICATIONS OF THE THEORY

A. General

In the present section we illustrate the theory by applying it to some explicit examples of two-dimensional waveguide structures with a refractive index depending on only one transverse coordinate, $n(\mathbf{r}) = n(x)$, which trivially allows the reduction of the problem to a scalar case, as has been shown above. Such potentials can be manufactored by modern dielectric layer deposition techniques. We proceed as follows: (i) First we solve the propagation of the scalar electric field $E(x,z)$ by calculating the eigenfunctions of the Helmholtz equation; (ii) then we compare this with the time-dependent solution of the optical Schrödinger equation for the same initial condition; (iii) finally we add the higher-order ''optical corrections'' to the Schrödinger solution to get a better approximation of the Helmholtz solution.

Before we carry out this procedure, let us first make a general observation. The propagation of an initial wave packet along the waveguide is simply given by the superposition

$$E(x,z) = \sum_\nu h_\nu \varphi_\nu^H(x) e^{ik_\nu z}, \tag{51}$$

where h_ν represent the expansion coefficients of the initial wave packet in the eigenfunctions $\{\varphi_\nu^H(x)\}$ of the Helmholtz equation

$$\nabla^2 E(x,z) + k_0^2 n^2(x) E(x,z) = 0. \tag{52}$$

Rewriting this equation as an eigenfunction problem

$$-\frac{\partial^2}{\partial x^2}\varphi_\nu^H(x) - k_0^2 n^2(x)\varphi_\nu^H(x) = E_\nu^H \varphi_\nu^H(x), \quad \text{with } E_\nu^H \equiv -k_\nu^2, \tag{53}$$

and comparing this with the time-independent form of the optical Schrödinger equation,

$$-\frac{\partial^2}{\partial x^2}\varphi_\nu^S(x) - k_0^2[n^2(x) - n_0^2]\varphi_\nu^S(x) = 2k_0^2 n_0^2 E_\nu^S \varphi_\nu^S(x), \tag{54}$$

we immediately see that the corresponding eigenfunctions $\{\varphi_\nu^H(x)\}$ and $\{\varphi_\nu^S(x)\}$ coincide, but belong to different eigenvalues, i.e.,

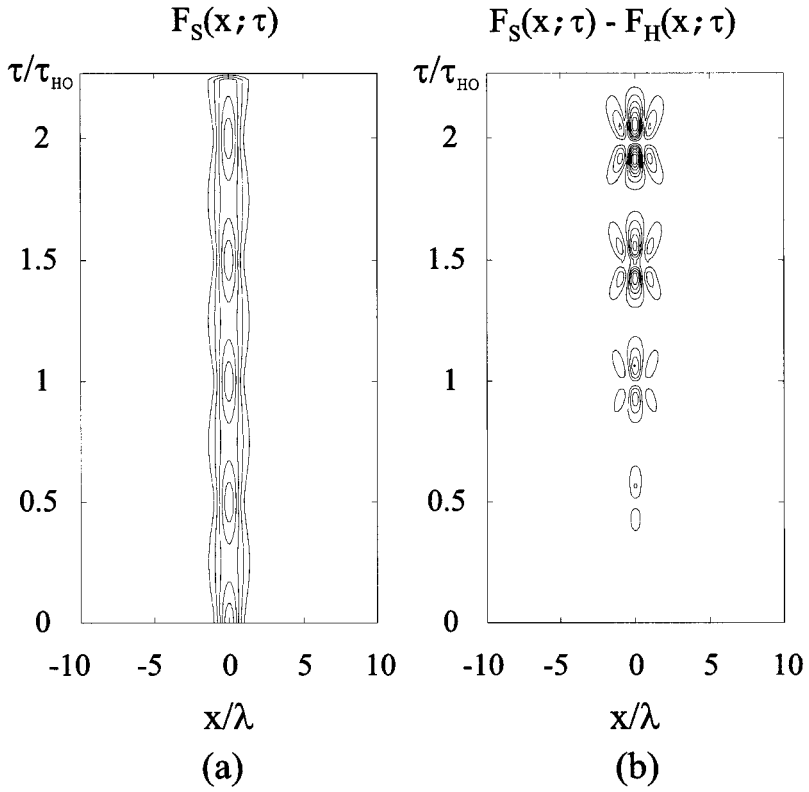

FIG. 2. Position distribution of the Schrödinger wave packet in the harmonic potential (a) and differences between the Schrödinger and the Helmholtz position distributions (b) for an initial Gaussian wave packet with $x_0=0$ and $\Delta x_0=0.8\Delta x_{\rm coh}$ (''squeezed vacuum''); in (b) the differences grow to a maximum of about 5%.

$$E_\nu^H=k_0^2n_0^2(2E_\nu^S-1) \quad \text{for} \quad \varphi_\nu^H(x)=\varphi_\nu^S(x). \tag{55}$$

B. The harmonic oscillator

One of the simplest possible examples to test our theory on is harmonic motion. We assume that the refractive index is a quadratic function of the transverse coordinate x,

$$n^2(\mathbf{r})=n_0^2(1-\kappa^2x^2). \tag{56}$$

In the zeroth-order paraxial approximation this leads to an optical Schrödinger equation with a harmonic potential

$$V_{\rm opt}(x)=\frac{1}{2}\kappa^2x^2 \tag{57}$$

and an oscillation frequency $\omega_{\rm HO}=\kappa/\hbar k_0n_0$. Obviously there is a restriction on the ''coupling strength'' κ, since $n^2(\mathbf{r})$ has to stay close to n_0^2, i.e., $\kappa^2x^2\ll1$, over the range of interest of x.

Figures 2 and 3 compare solutions for the time evolution of an initial Gaussian wave packet obtained from the Helmholtz and Schrödinger equations. For the light wave, the τ axis is proportional to the propagation direction z of the wave packet in the confining refractive index profile; for the corresponding fictitious Schrödinger particle $\tau=\hbar k_0n_0z$ has the meaning of the evolution time in the harmonic well. On the left-hand side the Schrödinger equation position distribution $F_S(x;\tau)=|\psi(x;\tau)|^2$ is illustrated by means of a contour plot, and it is compared with the solution $F_H(x,z)=|E(x,z)|^2$ of the Helmholtz equation on the right-hand side. We have taken $n_0=2$ and $\kappa^2=8\times10^{-3}/\lambda^2$ with $\lambda=2\pi/k_0$ being the wavelength of the carrier wave in vacuum; λ serves as a typical length scale of the problem. The interaction time τ increases to twice the harmonic oscillation time $\tau_{\rm HO}=2\pi/\omega_{\rm HO}$.

The differences between the Helmholtz and the Schrödinger solutions are illustrated for two initial conditions. In Fig. 2 we have chosen an initial Gaussian wave packet with its width Δx_0 being 80% of the width $\Delta x_{\rm coh}=1/\sqrt{2k_0n_0\kappa}$ of the ground state in the harmonic potential and an initial displacement $x_0/\lambda=0$; i.e., a ''squeezed vacuum'' in the quantum optics language. In Fig. 3 the initial width is equal to $\Delta x_{\rm coh}$ and $x_0/\lambda=-1$, i.e., a ''coherent state.'' In the first case the width of the Schrödinger position distribution, depicted in Fig. 2(a), undergoes a ''breathing'' evolution, which may be viewed as a built-in modulation of the paraxiality parameter $\Theta\approx\Delta k_\tau/k_0$ with Δk_τ being the k uncertainty in the wave packet during the evolution, $\Delta k_\tau=\sqrt{\langle k^2\rangle_\tau-\langle k\rangle_\tau^2}$. The ''butterfly corrections'' in Fig. 2(b) are seen to be relatively small, whenever the width Δx_τ of the wave packet is large and thus Δk_τ small; simultaneously they grow in time to a maximum value of about 5% in the picture. In the case in Fig. 3 the width of the wave packet remains constant, but the mean transverse k vector $\langle k\rangle_\tau$ oscillates in time; thus the discrepancies (which grow to 20%) are seen to be relatively small whenever the Schrödinger wave packet is near the turning points of the harmonic well, i.e., whenever $\langle k\rangle_\tau\approx0$.

Clearly wave packets with large initial displacement from the bottom of the harmonic well are less ''paraxial'' than wave packets with $x_0/\lambda\approx0$. This intuitive fact is also evident in the typical time scale $\tau_{\rm corr}\sim(E_{\nu_0}^S)^{-2}$ over which the lowest-order correction grows: for highly excited states the eigenvalue $E_{\nu_0}^S$ of the eigenfunction of maximal overlap with

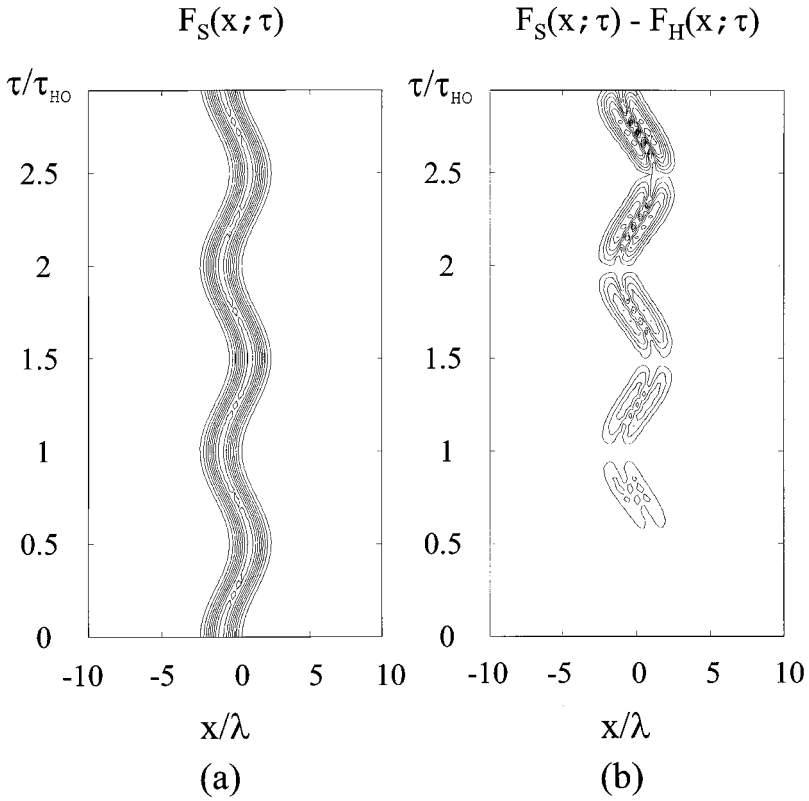

FIG. 3. Position distribution of the Schrödinger wave packet in the harmonic potential (a) and differences between the Schrödinger and the Helmholtz position distributions (b) for an initial Gaussian wave packet with $x_0=-1$ and $\Delta x_0=\Delta x_{\rm coh}$ (''coherent state''); in (b) the differences grow to a maximum of about 20%.

the initial state is larger; consequently the propagation time scale over which the paraxial approximation is reasonable becomes correspondingly shorter.

In Fig. 4 the first few order optical corrections are shown for the case of Fig. 2: In Fig. 4(a) the *difference* between the real parts of the slowly varying Schrödinger and Helmholtz solutions is compared with the first-order correction term; the difference between *these two* is then compared with the second-order correction in Fig. 4(b), the remaining difference with the next-order correction in Fig. 4(c), and so forth. We see that as long as the total interaction time is only a fraction of $\tau_{\rm corr}$ (in Fig. 4 we have $\tau_{\rm final}=1.5\tau_{\rm HO}=0.008\tau_{\rm corr}$), the first few terms of the above series expansion indeed yield an excellent approximation to the exact Helmholtz field. The fast oscillations turning up in Fig. 4(d) stem from numerical errors, as will be discussed briefly at the end of this section.

C. Tunneling (double-well potential)

Now we proceed to a more challenging and physically more interesting problem: What kind of optical corrections is one to expect for tunneling in a double-well potential? On the optical side this could be viewed as a simple model for an optical coupler for a paraxial light beam initially confined to one well [10].

A refractive index variation of the form

$$n^2(\mathbf{r})=n_0^2[1-d^4(x^2-a^2)^2] \tag{58}$$

leads to a double-well optical potential

$$V_{\rm opt}(x)=\frac{1}{2}d^4(x^2-a^2)^2, \tag{59}$$

again with the requirement for d that $d^4(x^2-a^2)^2\ll 1$ over the spatial interval of interest, which scales as a. Setting

$$d^4=\frac{(1-1/n_0^2)}{L^4\left[\frac{1}{4}-(a/L)^2\right]^2} \tag{60}$$

the refractive index takes on the vacuum value $n(\pm L/2)=1$ at the end points of the interval $[-L/2,L/2]$. For a given L, the depth and the location of the two potential wells are thus both fixed by the parameter a.

For our numerical illustrations we have chosen $L=5\lambda$ and $a/L=0.225$ throughout, and for reasons that will be explained below, instead of a Gaussian wave packet as initial condition, we choose a superposition of the two lowest eigenfunctions of the Schrödinger equation

$$\psi(x;\tau=0)=\frac{1}{\sqrt{2}}[\varphi_0(x)+\varphi_1(x)]. \tag{61}$$

In Fig. 5(a), where we have set $n_0=2.5$, the tunneling of the wave packet between the two potential wells is clearly visible. The time scale of the motion back and forth is determined by the inverse tunneling rate

$$\tau_{\rm tun}/\hbar=\frac{\pi}{(E_1^S-E_0^S)}, \tag{62}$$

which depends on the difference between the lowest two eigenenergies of the optical Schrödinger equation. So far this is only a simple Schrödinger evolution, but is it possible to get similar behavior in the corresponding paraxial Helmholtz problem? In other words, can the ''tunneling'' of the wave

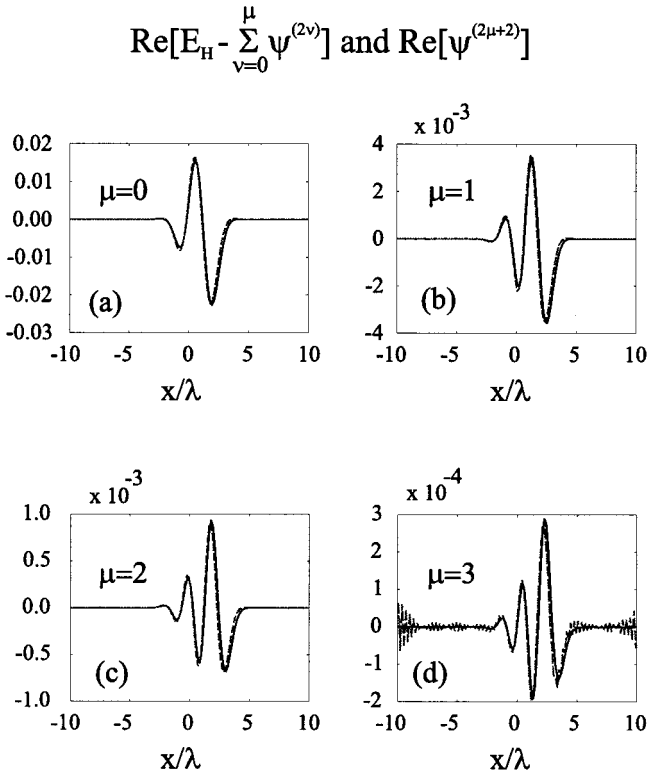

FIG. 4. The optical corrections for the same initial condition as in Fig. 2 after an evolution time $\tau_{\text{final}}=1.5\tau_{\text{HO}}=0.008\tau_{\text{corr}}$: (a) difference of the real parts of the slowly varying amplitude of the Helmholtz and the Schrödinger field (solid curve) and first-order correction (dashed curve), (b)–(d) remaining differences (solid curve) and the next higher-order correction (dashed curve).

packet from one graded-index well to the other in a coupler be calculated within the paraxial approximation?

It turns out that the relative size of two time scales is relevant for this question. For the paraxial description to be meaningful for times of the order of τ_{tun}, we require the typical time scale of the corrections $\tau_{\text{corr}}\sim(E_1^S)^{-2}$, introduced previously, to satisfy

$$\tau_{\text{corr}}>\tau_{\text{tun}}\,. \tag{63}$$

For the parameters chosen in Fig. 5 we have $\tau_{\text{tun}}/\tau_{\text{corr}}=0.40$. Therefore the discrepancies between the Schrödinger and the Helmholtz position distributions remain small, at most 5% in Fig. 5(b). As in the harmonic oscillator example the deviations are smallest around the turning points, where the transverse velocity and thus the paraxiality parameter are smallest.

The significance of the first few terms of the expansion is demonstrated in Fig. 6, where the *difference* between the real part $\text{Re}[\psi^{(0)}(x;z)]$ of the Schrödinger wave function and the real part of the slowly varying Helmholtz field amplitude $\text{Re}[E(x,z)\exp\{-ik_0n_0z\}]$ is plotted. From (a) to (d) an increasing number of correction terms is taken into account; the more correction terms are included, the later the onset of deviations. Note that even for the relatively long final propagation time $\tau_{\text{final}}=2.5\tau_{\text{tun}}(=1.01\tau_{\text{corr}})$ in Fig. 6, excellent agreement is reached when including terms up to order $\mu=3,4$.

Now we increase the refractive index value. This affects the ''mass'' $m_{\text{opt}}=k_0^2n_0^2$ of the fictitious Schrödinger particle and increases the difference between the Schrödinger eigenenergies $k_0^2n_0^2\{E_\nu^S\}$ and $\{E_\nu^H=-k_\nu^2\}$ belonging to the shared eigenfunctions $\{\varphi_\nu\}$. For $n_0=3$ the tunneling time scale τ_{tun} is now of the same order of magnitude as the correction time scale, $\tau_{\text{tun}}/\tau_{\text{corr}}\approx1$, in contrast to the previous case. This means that the tunneling to the other well may be described only poorly in the paraxial approximation. This is depicted in Fig. 7; as in Fig. 6 we plot the differences in the amplitude of the two fields, including terms up to order μ. The final time of $\tau_{\text{final}}=2.5\tau_{\text{tun}}$ now corresponds to $\tau_{\text{final}}=2.8\tau_{\text{corr}}$ and the truncation is seen to fail when $\tau\approx\tau_{\text{corr}}$; note the scaling factor at the height of about 50 as compared with Fig. 6.

Since the Schrödinger and the Helmholtz eigenfunctions coincide, as explained above, any initial condition involving a *finite* number of eigenstates, such as the example in the preceding section, will again be restored exactly after some finite time of evolution. Therefore for the initial condition discussed above, we have a third time scale of relevance, the ''dephasing'' time scale, which is proportional to the inverse of the difference between the lowest level splittings Δ_S and Δ_H of the Schrödinger and Helmholtz eigenvalues (in appropriate dimensionless units for comparison),

$$\tau_{\text{dephase}}/\hbar=\frac{\pi}{|\Delta_S-\Delta_H|}, \tag{64}$$

with

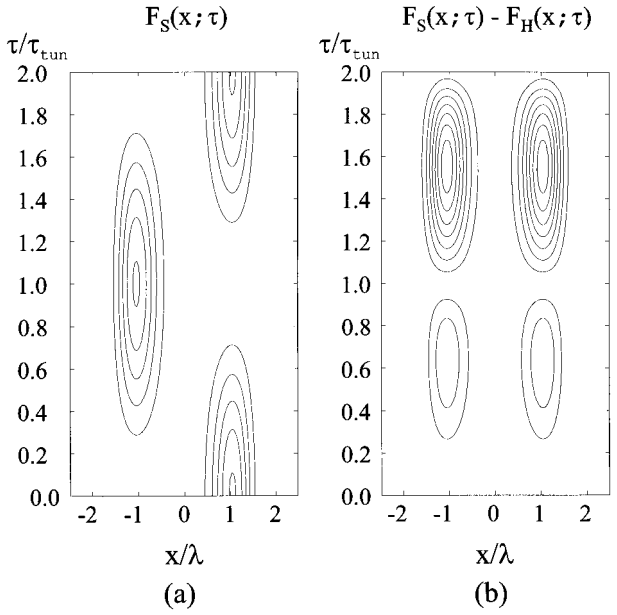

FIG. 5. Position distribution of the Schrödinger wave packet in the double-well potential with $n_0=2.5$, $a=0.225L$: (a) and differences between the Schrödinger and the Helmholtz position distributions (b) for $\tau_{\text{final}}=2\tau_{\text{tun}}=0.8\tau_{\text{corr}}$; in (b) the differences grow to a maximum of about 5%.

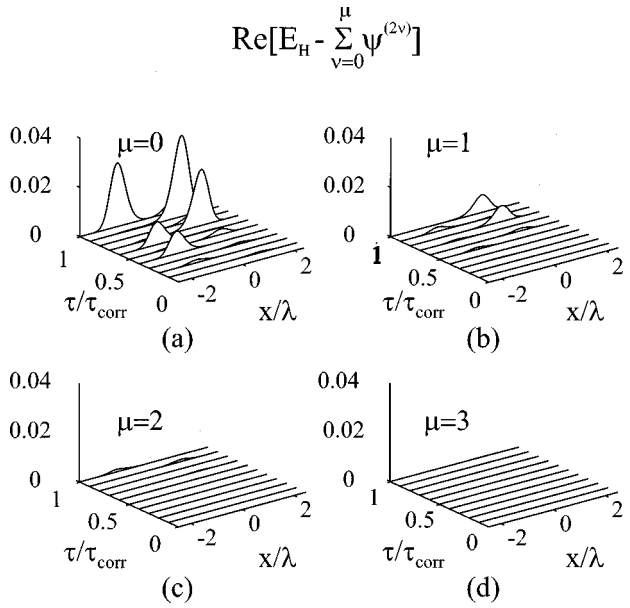

FIG. 6. The optical corrections for the same parameters as in Fig. 6 and $\tau_{\text{final}}=2.5\tau_{\text{tun}}=1.01\tau_{\text{corr}}$; from (a) to (d): difference between the real part of the Helmholtz amplitude and the Schrödinger amplitude with added corrections up to order μ.

$$\Delta_S = E_1^S - E_0^S,$$

$$\Delta_H = \frac{\sqrt{-E_1^H} - \sqrt{-E_0^H}}{n_0 k_0}. \tag{65}$$

Since Δ_S and Δ_H determine the individual tunneling rates, for $\tau_{\text{dephase}}|\Delta_S - \Delta_H|/\hbar = \pi$ the two fields have run completely out of phase, as shown in Fig. 8. After having tunneled back and forth about 30 times, the fields are localized in opposite wells for the Schrödinger and Helmholtz cases.

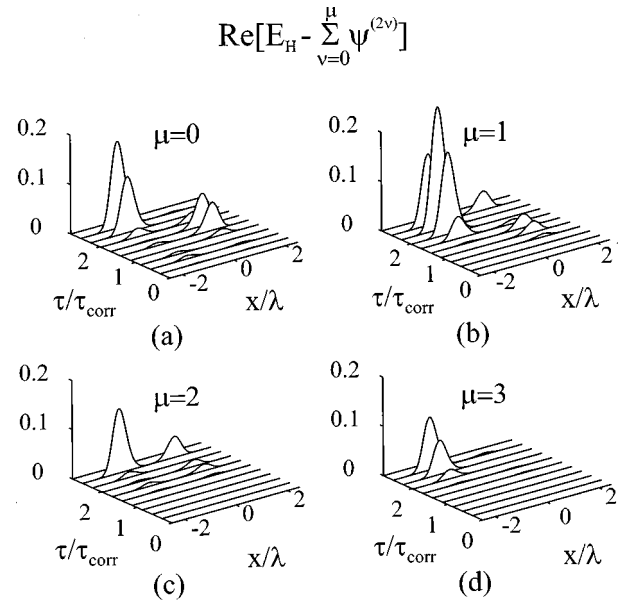

FIG. 7. The optical corrections for a potential with $n_0=3$ and $\tau_{\text{final}}=2.5\tau_{\text{tun}}=2.8\tau_{\text{corr}}$; from (a) to (d): difference between the real part of Helmholtz amplitude and the Schrödinger amplitude with added corrections up to order μ.

Here $\tau_{\text{final}}=\tau_{\text{dephase}}=34\tau_{\text{tun}}=133.5\tau_{\text{corr}}$, and a truncation of the expansion series would fail completely.

The convergence range of the expansion Eq. (40) is not clear, it is certainly not uniform in time. We stress that reliable results are only guaranteed for $\tau_{\text{final}} \ll \tau_{\text{corr}}$ and recommend the following practical approach for estimating the validity of the truncation of the infinite series to a given finite number of terms: Higher-order contributions subtract from the norm of the total wave function, canceling the overwhelming power law growth of the low-order correction terms. Checking the normalization of the total wave function including terms up to order μ therefore has proven to be an excellent way to test the validity of truncation at a given order μ. For the situation shown in Fig. 6, for instance, the normalization of the wave function is 1.71, 1.15, 0.97, 0.99, . . . when including corrections up to $\mu=0,1,2,3,\ldots$. Furthermore we want to mention that, whenever possible, we have also solved the dynamics by means of another independent method, e.g., fast Fourier transform, in order to verify the numerical accuracy.

Finally a word on choosing a superposition of the two lowest eigenstates instead of a Gaussian located in one of the potential wells as initial condition. It turns out that one encounters numerical problems, which become more serious the more coefficients have to be taken into account when expanding the Gaussian in the eigenbasis. The reason for this is that on a finite grid high-order eigenfunctions are not represented properly. Since the evaluation of the wave function up to order μ involves a summation over the eigenindex n which involves terms with increasingly fast oscillations of their phases, such errors are of critical importance.

V. DISCUSSION

In this paper we have compared three different calculations, viz., the Schrödinger propagation, which is identical with the paraxial propagation of light, the systematic perturbation corrections to this lowest-order result, and finally the exact electromagnetic propagation described by the Helmholtz equation. The perturbation expansion is of the singular kind, which produces unbounded corrections valid over a limited interval only. Such terms, however, suffice for our purpose to investigate the validity range of the paraxial approximation and its corrections. The correct result can always be obtained from the Helmholtz equation.

The theory is applied to simple models of wave propagation in graded-index media. In Figs. 2 and 3 we can see how the paraxial evaluation of the harmonic motion becomes exact whenever the wave packet focuses in the forward direction, making the transverse momentum components small. The illustrations also show the steady growth of the error over a few oscillation periods; this growth is illustrated in Fig. 4.

We also apply the general theory to the case of propagation in two coupled potential wells. Here the wave packet is expected to tunnel between the wells, and we investigate the differences between Schrödinger theory tunneling and ''tunneling'' of electromagnetic waves. The results are reported in Figs. 5–7, where the main conclusions from the simple harmonic case are verified. An interesting observation is that the electromagnetic wave and the massive particle develop a

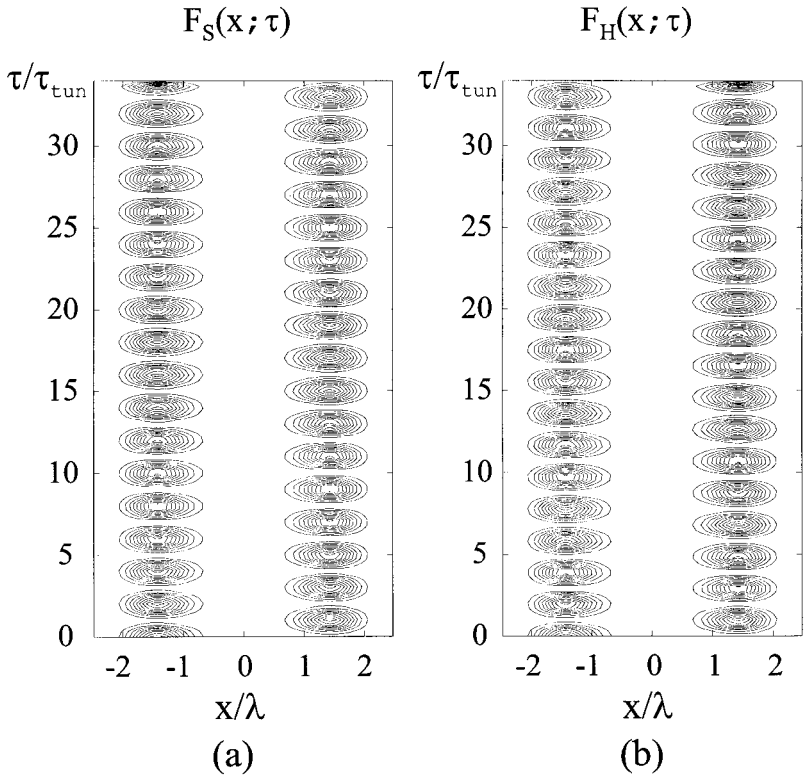

FIG. 8. Position distribution of the Schrödinger (a) and the Helmholtz (b) wave packet in the double-well potential with $n_0=1.5, a=0.3L$ after the time $\tau_{\text{final}}=\tau_{\text{dephase}}=34\tau_{\text{tun}}=133.5\tau_{\text{corr}}$. After having tunneled about 30 times, the two fields have run out of phase and are localized in opposite wells.

phase difference when they propagate; this is shown in Fig. 8. The failure of the straightforward perturbation approach with increasing time is derived directly from this phase difference. When the theory tries to approximate this in a perturbative manner, it is bound to produce terms growing as some power of the propagation time. Only a uniform expansion like the multiple-time-scale one would be able to treat such effects appropriately. We have not regarded it expedient to pursue such an approach in the present paper.

Our main objective here has been to extract those mathematical terms which make the Schrödinger evolution start to differ from electromagnetic wave propagation. It is, moreover, of some interest to compare the physical pictures offered by the two theories. In this spirit let us conclude by looking at the scaling relations of the parameters occurring in the Schrödinger interpretation of the electromagnetic wave and its wave characteristics. This can be done using some elementary dimensional arguments in the following manner.

As we have seen, the lowest order paraxial equation (3) can be written

$$i\frac{\partial}{\partial z}\psi=-\frac{1}{2k_0}\nabla_\perp^2\psi+V_{\text{opt}}\psi,$$

where for notational simplicity we have assumed $n_0=1$. Introducing a properly scaled time variable $t=z/c$, we find the effective Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi=-\frac{\hbar^2}{2m_{\text{eff}}}\nabla_\perp^2\psi+V_{\text{eff}}\psi,$$

where the role of the mass is played by the parameter

$$m_{\text{eff}}=\frac{\hbar k_0}{c}, \tag{66}$$

as we might expect on dimensional grounds.

For the massive particle case, we expect a heavy particle to display nearly classical behavior. It is now of interest to ask what this requirement implies for the electromagnetic case in the paraxial approximation. If we assume that the potential function is characterized by a length scale L, we can define a characteristic time scale L/c. If we inject an initial wave packet of transverse width Δx, we can neglect the dispersion as long as

$$\Delta x\gg\frac{L\Delta p}{mc},$$

where Δp is the momentum width of the wave packet. From this relation and the uncertainty relation, we obtain the condition on the mass

$$m_{\text{eff}}\gg\frac{L\hbar}{c\Delta x^2}.$$

When the effective mass (66) is inserted into this condition we find the relation

$$\frac{\Delta x^2}{\lambda L}\gg 1, \tag{67}$$

which is the Fresnel condition for ray behavior; this is well known to correspond to the classical limit of wave mechanics. We can combine this with the condition for the validity

of a semiclassical description, namely, that the potential structure changes slowly over the extent of the transverse de Broglie wavelength

$$\lambda_\perp = \frac{\lambda}{\Theta}.$$

With our definition of the characteristic length L we can write this as

$$\lambda \ll \Theta L.$$

Combined with Eq. (67) this gives the condition

$$\left(\frac{\Delta x}{\lambda_\perp}\right)^2 \gg \Theta.$$

This states the well known fact that a wave packet can be localized only to within a region much larger than its average wavelength; here it is applied to the transverse direction, which in the paraxial approximation involves the expansion angle Θ. When this goes to zero, the Fresnel condition (67) is sufficient to localize the particle wave packet.

Finally a few words on possible extensions of the present work. In the future it might be worthwhile to address problems that involve the true vector character of the electromagnetic field: however, if the three spatial components do not decouple, one can no longer associate one single optical Schrödinger equation with the paraxial wave packet; it does not even seem possible to generalize this to a (nonlinear) Schrödinger equation for coupled particles, since the equation is not homogeneous in the wave functions. Moreover, in the present work we have only dealt with monochromatic fields, that is, we have ignored temporal effects.

ACKNOWLEDGMENTS

M.M. was supported by the *APART* Program of the Österreichische Akademie der Wissenschaften and wishes to thank H. Ritsch for critical remarks and suggestions.

APPENDIX

1. The general vector case for monochromatic fields

For monochromatic fields the electric field vector is given by $\mathbf{E}(\mathbf{r},t)=\mathbf{E}(\mathbf{r})\exp(-i\omega t)$, with $\mathbf{E}(\mathbf{r})$ satisfying

$$k_0^2 n^2(\mathbf{r})\mathbf{E}(\mathbf{r}) = \nabla\times\nabla\times\mathbf{E}. \tag{A1}$$

Pulling out an appropriate phase factor

$$\mathbf{E}(\mathbf{r}) = \mathcal{E}(\mathbf{r}) e^{i\varphi(\mathbf{r})} \tag{A2}$$

this is equivalent to the envelope equation

$$k_0^2 n^2(\mathbf{r})\mathcal{E}(\mathbf{r}) = [\nabla_\varphi(\nabla_\varphi\cdot\mathcal{E}(\mathbf{r})) - \nabla_\varphi^2\mathcal{E}(\mathbf{r})], \tag{A3}$$

where $\nabla_\varphi = \nabla + i(\nabla\varphi)$.

Now we transform to scaled variables, which for convenience we write down again,

$$\begin{aligned}\tilde{\mathbf{r}}_\perp &= \Theta k_0 n_0 \mathbf{r}_\perp,\\ \tilde{z} &= \Theta^2 k_0 n_0 z,\\ \tilde{t} &= \Theta^2 \omega_0 t,\end{aligned} \tag{A4}$$

and collect all terms of the same order in Θ in the differential operator on the right-hand side of Eq. (A3). With the definitions

$$\Gamma_{\tilde{x}_i} = \frac{\partial}{\partial\tilde{x}_i}[\Theta^2\tilde{\varphi}(\tilde{\mathbf{r}})] \equiv \begin{cases} \dfrac{\Theta}{k_0 n_0}\dfrac{\partial\varphi(\mathbf{r})}{\partial x_i} & \text{for } x_i\in\{x,y\} \\ \dfrac{1}{k_0 n_0}\dfrac{\partial\varphi(\mathbf{r})}{\partial x_i} & \text{for } x_i = z \end{cases} \tag{A5}$$

we get the three coupled equations for the three Cartesian components

$$\begin{aligned}0 = &-\Theta^4\frac{\partial^2\mathcal{E}_x}{\partial\tilde{z}^2} + \Theta^3\frac{\partial^2\mathcal{E}_z}{\partial\tilde{x}\partial\tilde{z}} + \Theta^2\left(-2i\Gamma_{\tilde{z}}\frac{\partial\mathcal{E}_x}{\partial\tilde{z}} - i\frac{\partial\Gamma_{\tilde{z}}}{\partial\tilde{z}}\mathcal{E}_x\right.\\ &\left. - \frac{\partial^2\mathcal{E}_x}{\partial\tilde{y}\partial\tilde{y}} + \frac{\partial^2\mathcal{E}_y}{\partial\tilde{x}\partial\tilde{y}}\right) + i\Theta^1\left(\Gamma_{\tilde{x}}\frac{\partial\mathcal{E}_z}{\partial\tilde{z}} + \Gamma_{\tilde{z}}\frac{\partial\mathcal{E}_z}{\partial\tilde{x}} + \frac{\partial\Gamma_{\tilde{x}}}{\partial\tilde{z}}\mathcal{E}_z\right)\\ &+ \Theta^0\left(\Gamma_{\tilde{z}}^2\mathcal{E}_x - 2i\Gamma_{\tilde{y}}\frac{\partial\mathcal{E}_x}{\partial\tilde{y}} - i\frac{\partial\Gamma_{\tilde{y}}}{\partial\tilde{y}}\mathcal{E}_x + \Gamma_{\tilde{x}}\frac{\partial\mathcal{E}_y}{\partial\tilde{y}} + i\frac{\partial\Gamma_{\tilde{x}}}{\partial\tilde{y}}\mathcal{E}_y\right.\\ &\left. + \Gamma_{\tilde{y}}\frac{\partial\mathcal{E}_y}{\partial\tilde{x}}\right) - \Theta^{-1}\Gamma_{\tilde{x}}\Gamma_{\tilde{z}}\mathcal{E}_z - \Theta^{-2}(\Gamma_{\tilde{x}}\Gamma_{\tilde{y}}\mathcal{E}_y - \Gamma_{\tilde{y}}^2\mathcal{E}_x),\end{aligned}$$

$$0 = [\cdots]_{x\leftrightarrow y}, \tag{A6}$$

$$\begin{aligned}0 = &\Theta^3\left(\frac{\partial^2\mathcal{E}_x}{\partial\tilde{x}\partial\tilde{z}} + \frac{\partial^2\mathcal{E}_y}{\partial\tilde{y}\partial\tilde{z}}\right) + \Theta^2\left(\frac{\partial^2\mathcal{E}_z}{\partial\tilde{x}^2} + \frac{\partial^2\mathcal{E}_z}{\partial\tilde{y}^2}\right) + i\Theta^1\left[\Gamma_{\tilde{x}}\frac{\partial\mathcal{E}_x}{\partial\tilde{z}}\right.\\ &\left. + \Gamma_{\tilde{y}}\frac{\partial\mathcal{E}_y}{\partial\tilde{z}} + \Gamma_{\tilde{z}}\left(\frac{\partial\mathcal{E}_x}{\partial\tilde{x}} + \frac{\partial\mathcal{E}_y}{\partial\tilde{y}}\right) + \frac{\partial\Gamma_{\tilde{x}}}{\partial\tilde{z}}\mathcal{E}_x + \frac{\partial\Gamma_{\tilde{y}}}{\partial\tilde{z}}\mathcal{E}_y\right]\\ &- i\Theta^0\left[2\Gamma_{\tilde{x}}\frac{\partial\mathcal{E}_z}{\partial\tilde{x}} + 2\Gamma_{\tilde{y}}\frac{\partial\mathcal{E}_z}{\partial\tilde{y}} + \left(\frac{\partial\Gamma_{\tilde{x}}}{\partial\tilde{x}} + \frac{\partial\Gamma_{\tilde{y}}}{\partial\tilde{y}}\right)\mathcal{E}_z\right]\\ &- \Theta^{-1}\Gamma_{\tilde{z}}(\Gamma_{\tilde{x}}\mathcal{E}_x + \Gamma_{\tilde{y}}\mathcal{E}_y) + \Theta^{-2}(\Gamma_{\tilde{x}} + \Gamma_{\tilde{y}})\mathcal{E}_z.\end{aligned}$$

When setting

$$\nabla_\varphi\cdot\mathcal{E} = 0, \tag{A7}$$

the complicated expressions simplify considerably; all odd-order terms vanish and all components decouple, each satisfying the same scalar equation (given in the next section below). Writing this condition in scaled variables,

$$\Theta^3\frac{\partial\mathcal{E}_z}{\partial\tilde{z}}+\Theta^2\left(\frac{\partial\mathcal{E}_x}{\partial\tilde{x}}+\frac{\partial\mathcal{E}_y}{\partial\tilde{y}}\right)+i\Theta\Gamma_{\tilde{z}}\mathcal{E}_z+i(\Gamma_{\tilde{x}}\mathcal{E}_x+\Gamma_{\tilde{y}}\mathcal{E}_y)=0, \tag{A8}$$

makes it obvious that the longitudinal component $\mathcal{E}_z$, which leads to a coupling between all components, is also responsible for the odd-order terms in Θ.

2. An alternative form of the expansion for the scalar wave equation

In this section we will give an alternative derivation of Eqs. (41) and (42). Our starting point is the wave equation for the slowly varying amplitude $\psi(\mathbf{r},t)=E(\mathbf{r},t)e^{-i\varphi(\mathbf{r})+i\omega t}$ for a solution of the scalar wave equation

$$-\frac{n^2(\mathbf{r})}{c^2}\left(\frac{\partial}{\partial t}-i\omega_0\right)^2\psi(\mathbf{r},t)=-\nabla_\varphi^2\psi(\mathbf{r},t) \tag{A9}$$

or equivalently, by formally taking the square root on both sides [9],

$$i\frac{n(\mathbf{r})}{c}\left(\frac{\partial}{\partial t}-i\omega_0\right)\psi(\mathbf{r},t)=\sqrt{-\nabla_\varphi^2}\psi(\mathbf{r},t). \tag{A10}$$

Here we have generalized from a strictly monochromatic wave to a quasimonochromatic one. In scaled variables we get

$$i\frac{\partial}{\partial\tilde{t}}\psi(\tilde{\mathbf{r}},\tilde{t})=\frac{1}{\Theta^2}\left[\frac{n_0}{n(\tilde{\mathbf{r}})}\sqrt{\Theta^4\Delta_4+\Theta^2\Delta_2+\Delta_0+\Theta^{-2}\Delta_{-2}}-1\right]\psi(\tilde{\mathbf{r}},\tilde{t}), \tag{A11}$$

with the definitions

$$\Delta_4=-\frac{\partial^2}{\partial\tilde{z}^2},$$

$$\Delta_2=-\tilde{\nabla}_\perp^2-i\frac{\partial\Gamma_{\tilde{z}}}{\partial\tilde{z}}-2i\Gamma_{\tilde{z}}\frac{\partial}{\partial\tilde{z}}, \tag{A12}$$

$$\Delta_0=\Gamma_{\tilde{z}}^2-i\left(\frac{\partial\Gamma_{\tilde{x}}}{\partial\tilde{x}}+\frac{\partial\Gamma_{\tilde{y}}}{\partial\tilde{y}}\right)-2i\left(\Gamma_{\tilde{x}}\frac{\partial}{\partial\tilde{x}}+\Gamma_{\tilde{y}}\frac{\partial}{\partial\tilde{y}}\right),$$

$$\Delta_{-2}=\Gamma_{\tilde{x}}^2+\Gamma_{\tilde{y}}^2,$$

Eq. (A11) is still exact for arbitrary $\varphi(\mathbf{r})$. It contains all the information of spatial and temporal effects on the propagating wave packet (cf. right- and left-hand side, respectively), except for the vector aspects of the field discussed in the preceding section.

Assuming a quadratic scaling (38) in Θ of the smoothness of the refractive index profile, we may rewrite

$$\frac{n_0}{n(\tilde{\mathbf{r}})}=\frac{1}{\sqrt{n(\tilde{\mathbf{r}})^2/n_0^2}}=\frac{1}{\sqrt{1+\Theta^2\delta n(\tilde{\mathbf{r}})}}, \tag{A13}$$

with $\delta n(\tilde{\mathbf{r}})=O(1)$. Now one has to expand the whole expression inside the square brackets in Eq. (A11) in a power series of Θ, which may be cumbersome for complicated transverse and/or longitudinal refractive index profiles and appropriate exponents $\varphi(\mathbf{r})$, which suppress the divergent terms $\sim\Delta_{-2}$.

From here on we specialize again to the case of monochromatic fields treated in the main text [with $(\partial/\partial\tilde{t})\psi=0$ and $\varphi=k_0n_0z$]. We get

$$\Gamma_{\tilde{x}}^2=\Gamma_{\tilde{y}}^2=0,\quad\Gamma_{\tilde{z}}^2=1 \tag{A14}$$

and thus

$$\Delta_4=\frac{\partial^2}{\partial\tilde{z}^2},$$

$$\Delta_2=2H^{(0)}+\delta\tilde{n}, \tag{A15}$$

$$\Delta_0=1,$$

$$\Delta_{-2}=0,$$

where the operator $H^{(0)}$ is defined as

$$H^{(0)}=-i\frac{\partial}{\partial\tilde{z}}-\frac{1}{2}\tilde{\nabla}_\perp^2. \tag{A16}$$

Now we expand the operator in the square brackets on the right-hand side of Eq. (A11) and the wave function in a power series of Θ,

$$\psi=\sum_{\mu=0}^{\infty}\Theta^{2\mu}\psi^{(2\mu)},$$

$$[\ldots]=\sum_{\mu=0}^{\infty}\Theta^{2\mu}T^{(2\mu)}. \tag{A17}$$

The contributions satisfy

$$0=\sum_{\nu=0}^{\mu}T^{(2\mu-2\nu)}\psi^{(2\nu)}. \tag{A18}$$

In our case the lowest orders are found to be

$$T^{(0)}=H^{(0)}-\frac{\delta n}{2},$$

$$T^{(2)}=-\frac{1}{2}\frac{\partial^2}{\partial\tilde{z}^2}-\delta n T^{(0)}-\frac{1}{2}(T^{(0)})^2, \qquad \text{(A19)}$$

$$T^{(4)}=-T^{(2)}T^{(0)}+\frac{\delta n}{2}\frac{\partial^2}{\partial\tilde{z}^2}+\delta n^2(T^{(0)})^2.$$

In principle, for nonconstant δn one has to be careful with the operator ordering while expanding the square root. But since in Eq. (A11) δn is to the left of all differential operators, it has to be treated as a constant and factored out to the left in all terms of the above expressions. This and the fact that $\partial^2/\partial\tilde{z}^2$ and $H^{(0)}$ commute, imply that in the present simple case one can expand the square root as if dealing with c-numbers.

This expansion is equivalent to Eqs. (41). We verify this by inserting Eq. (A19) into Eq. (A17),

$$0=T^{(0)}\psi^{(0)},$$

$$0=T^{(0)}\psi^{(2)}+T^{(2)}\psi^{(0)}=T^{(0)}\psi^{(2)}-\frac{1}{2}\frac{\partial^2}{\partial\tilde{z}^2}\psi^{(0)}, \qquad \text{(A20)}$$

$$0=T^{(0)}\psi^{(4)}+T^{(2)}\psi^{(2)}+T^{(4)}\psi^{(0)}$$

$$=\cdots=T^{(0)}\psi^{(4)}-\frac{1}{2}\frac{\partial^2}{\partial\tilde{z}^2}\psi^{(2)}.$$

Here for every next higher order we have made iterative use of the lower-order relations (treating δn again as a c-number).

[1] A. Yariv, *Optical Electronics* (Saunders College Publishers, Philadelphia, 1991).

[2] A. W. Snyder and J. D. Love, *Waveguide Theory* (Chapman and Hall, London, 1991).

[3] P. W. Milonni, in *Coherence Phenomena in Atoms and Molecules in Laser Fields*, edited by A. D. Bandrauk and S. C. Wallace (Plenum, New York, 1992), p. 45.

[4] Ch. Bordé, in *Fundamental Systems in Quantum Optics*, Lecture Notes of the Les Houches Summer School Session LIII, edited by J. Dalibard, J.-M. Raimond, and J. Zinn-Justin (North-Holland, Amsterdam), p. 287.

[5] D. Marcuse, *Light Transmission Optics* (Van Nostrand Reinhold, New York, 1982).

[6] M. Lax, W. H. Louisell, and W. B. McKnight, Phys. Rev. A **11**, 1365 (1975).

[7] S. W. Koch and E. M. Wright, Phys. Rev. A **35**, 2542 (1987).

[8] I. Bialynicki-Birula, in *Coherence and Quantum Optics VII*, Proceedings of the Seventh Rochester Conference on Coherence and Qauntum Optics, edited by J. H. Eberly, L. Mandel, and E. Wolf (Plenum Press, New York, 1996), p. 313.

[9] I. H. Deutsch and J. C. Garrison, Phys. Rev. A **43**, 2498 (1991).

[10] P.-A. Bélanger, *Optical Fiber Theory* (World Scientific, Singapore, 1993).

[11] C. S. Adams, M. Sigel, and J. Mlynek, Phys. Rep. **240**, 144 (1994).

[12] M. Ben Dahan, E. Peik, J. Reichel, Y. Castin, and C. Salomon, Phys. Rev. Lett. **76**, 4508 (1996).

[13] S. R. Wilkinson, C. F. Bharucha, K. W. Madison, Q. Niu, and M. G. Raizen, Phys. Rev. Lett. **76**, 4512 (1996).

[14] M. K. Oberthaler, R. Abfalterer, S. Bernet, J. Schmiedmayer, and Z. Zeilinger, Phys. Rev. Lett. **77**, 4980 (1996).

[15] T. Takenaka, M. Yokota, and O. Fukumitsu, J. Opt. Soc. Am. B **2**, 826 (1985); S. Nemoto, Appl. Opt. **29**, 1940 (1990).

[16] F. Collino and P. Joly, SIAM (Soc. Ind. Appl. Math.) J. Sci. Comput. **16**, 1019 (1995).

[17] S. Solimeno, B. Crosignani, and P. DiPorto, *Guiding, Diffraction, and Confinement of Optical Radiation* (Academic Press, Orlando, 1986).

[18] P. Yeh, *Optical Waves in Layered Media* (Wiley, New York, 1988).

[19] L. D. Landau, E. M. Lifschitz, and L. P. Pitaewskii, *Electrodynamics of Continuous Media*, 2nd ed. (Pergamon Press, Oxford, 1984), p. 304.

[20] Yu. A. Kravtsov and Yu. I. Orlov, *Geometrical Optics of Inhomogeneous Media* (Springer, Berlin, 1990).

[21] J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics* (Springer-Verlag, New York, 1981).

[22] S. Stenholm and A. Bambini, IEEE J. Quantum Electron. **QE-17**, 1363 (1981).

[23] J. Mizerski, S. Stenholm, and W. Miklaszewski, J. Opt. Soc. Am. B **5**, 1600 (1988).