Accurate solution of the Helmholtz equation by Lanczos orthogonalization for media with loss or gain

R. P. Ratowsky, J. A. Fleck, Jr., and M. D. Feit

Lawrence Livermore National Laboratory, Livermore, California 94550

Received September 4, 1991

The numerical scheme for solving the Helmholtz equation, based on the Lanczos orthogonalization scheme, is generalized so that it can be applied to media with space-dependent absorption or gain profiles.

The Helmholtz equation plays a central role in the description of propagation phenomena in optics and acoustics. The paraxial approximation to the Helmholtz equation, also known as the paraxial wave equation, has long been the instrument of choice for the performance of calculations because it is amenable to solution by accurate marching techniques. The generation of an accurate solution to the unapproximated Helmholtz equation by marching, on the other hand, requires the accurate evaluation of a square-root operator applied to some initial field. Ideally, this evaluation could be performed in an appropriate diagonal representation having the same dimension as the number of points on the computational grid, but for most applications such a procedure would be impractical.

By using an orthogonalization procedure developed by Lanczos, however, one can generate a low-dimensional representation, valid over a sufficiently short propagation step, which accurately diagonalizes the square-root operator. This method is a generalization of earlier Lanczos-based methods for solving the time-dependent Schrödinger equation. An improved paraxial approximation to the Helmholtz equation can be derived from a pseudo-differential equation involving the square-root operator by approximating the square-root operator. The resulting wide-angle parabolic equation is, however, valid only for angular components up to 45 deg. The Lanczos solution, on the other hand, is valid for angles up to and including 90 deg. If either initial conditions or sharp changes in refractive index generate Fourier components of the field that do not correspond to forward-propagation angles, such components decay exponentially in the Lanczos scheme.

A shortcoming of the Lanczos propagation scheme as previously formulated for solving both the paraxial wave and the Helmholtz equations is that it is restricted to Hermitian operators, which prevents the use of imaginary refractive indices that represent the effects of gain or loss. This restriction excludes a wide class of interesting propagation applications. Even when loss is not explicitly included in the problem, it is customary to include loss in a border region along the grid boundary to prevent power from being reflected from the boundaries back into the interior of the computational grid. Even these losses cannot be included in a self-consistent way. One can apply a spatial filter to the field along the boundaries after each propagation step, but this may have subtle effects on long-range propagation. For example, we have found that a beam placed over one arm of a directional coupler will not necessarily couple to the other arm. Instead the entire field for certain input beams can relax to the lowest-order symmetrical mode of the coupler.

In this Letter we report a generalization of the standard Lanczos orthogonalization procedure to complex non-Hermitian operators. This generalization makes it possible to obtain accurate solutions to the Helmholtz equation with complex refractive-index distributions by the Lanczos orthogonalization procedure.

We begin by writing the scalar Helmholtz equation for a field variable \( \Psi \):

\[
\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} + \frac{n^2(x, y)\omega^2}{c^2} \Psi = 0, \tag{1}
\]

where \( n(x, y) = n'(x, y) [1 + i\delta(x, y)] \) is a complex refractive index and, for the present, independent of \( z \). The latter condition can be later relaxed as long as \( n \) depends weakly on \( z \). If we express the field variable in the form \( \Psi(x, y, z) = \exp(ikz)\phi(x, y, z) \), where \( k = \omega n_0/c \) and \( n_0 \) is a real reference refractive index, Eq. (1) becomes

\[
\frac{\partial^2 \phi}{\partial z^2} + 2ik \frac{\partial \phi}{\partial z} + V_{\perp} \phi + k^2 \left[ \frac{n^2(x, y)}{n_0^2} - 1 \right] \phi = 0. \tag{2}
\]

If we define the operator \( H \) as

\[
H = \nabla_{\perp}^2 + k^2 \left[ \frac{n^2(x, y)}{n_0^2} - 1 \right], \tag{3}
\]

Eq. (2) can be written as

\[
\frac{\partial^2 \phi}{\partial z^2} + 2ik \frac{\partial \phi}{\partial z} + H \phi = 0. \tag{4}
\]

Equation (4) is satisfied by two independent solutions that correspond to waves propagating to the right and to the left. The rightward-propagating solution can be written formally as

\[
\phi(z) = \exp[-iz(k - k(1 + H/k^2)^{1/2})] \psi(0). \tag{5}
\]
We wish to evaluate Eq. (5) by introducing a low-dimensional diagonal representation of the operator $H$. To this end we introduce as a basis the $N$ Krylov vectors $\psi(0), H\psi(0), \ldots H^{N-1}\psi(0)$, where the components of the vectors are the function values on the computational grid and the second derivatives in $H$ are evaluated by expressing $\psi(0)$ as a finite Fourier series. These vectors are independent but not orthonormal. Since the operator $H$ is non-Hermitian, the standard Lanczos orthogonalization procedure is not appropriate for deriving an orthonormal set of vectors from the Krylov vectors. We therefore generalize the standard Lanczos procedure as follows.

We define a set of right column vectors $|q_n\rangle, |q_{n-1}\rangle, \ldots |q_0\rangle$ and a set of left row vectors $\langle q_n|, \langle q_{n-1}|, \ldots \langle q_0|$.

The left-hand side of Eq. (6) signifies an inner product between $|q_n\rangle$ and $|q_n\rangle$, and the components of $|q_n\rangle$ are in general not the complex conjugates of the components of $|q_n\rangle$, as they are when $H$ is Hermitian. A generalization of the standard Lanczos orthogonalization recursion relation that leads to a symmetric representation of $H$ is given by the following pair of recursion relations:

$$\beta_n|q_{n+1}\rangle = H|q_n\rangle - \alpha_n|q_n\rangle - \beta_{n-1}|q_{n-1}\rangle,$$

(7a)

$$\beta_n|q_{n+1}\rangle = \langle q_n|H - \alpha_n\langle q_n| - \beta_{n-1}\langle q_{n-1}|,$$

(7b)

where $|q_0\rangle = \psi(0), \langle q_0| = \psi^*(0), \alpha_n = \langle q_n|H|q_n\rangle, \beta_n = |q_n\rangle H|q_n\rangle = \langle q_n^\dagger|H|q_n\rangle$. To compute $|q_n|H$ in Eq. (7b), one makes use of the relation $|q_n^\dagger|H = (H^\dagger|q_n^\dagger|)^\dagger$, where $H^\dagger$ represents the Hermitian conjugate of the operator defined in Eq. (3) and the remaining daggers signify the complex-conjugate transpose of the indicated vector.

Equations (7a) and (7b) determine $|q_{n+1}\rangle$ and $|q_{n+1}\rangle$ to within a normalization constant. If one takes the inner product of the left-hand sides of Eqs. (7a) and (7b) and equates the result to the inner product of the right-hand sides, one can evaluate $\beta_n^2$ since $|q_{n+1}^\dagger|q_{n+1}\rangle = 1$. Thus $\beta_n$ is determined to within an arbitrary choice of sign. The vectors $|q_{n+1}\rangle$ and $|q_{n+1}\rangle$ can then be completely determined from the recursion relations (7a) and (7b).

The matrix elements $\langle q_n|H|q_{n}\rangle$ form an $N$-dimensional tridiagonal complex symmetric representation of the operator $H$. The matrix, which we call $H_N$, can be reduced to diagonal form by the operation $\beta^r = \text{diag}(\beta_0, \beta_1, \ldots, \beta_{N-1}) = U H_N U^{-1}$, where $U U^\dagger = 1$. Here $U = (u_0, u_1, \ldots, u_{N-1})$, where the $u_n$ are the eigenvectors of $H$ and $U^{-1}$ is the transpose (without conjugation) of $U$. In this representation Eq. (5) can be evaluated by using the relation

$$\psi(x) = U^{-1} \exp[-iz(k - k(1 + \beta/k^2)]U\psi(0).$$

(8)

To test the solution method we consider the complex truncated quadratic index profile,

$$n^2 = n_1^2(1 - iv)^2 \left[ 1 - 2\Delta \left( \frac{x}{x_0} \right)^2 \right], \quad x < x_0,$$

$$n^2 = n_1^2(1 - 2\Delta), \quad x \geq x_0,$$

(9)

with $n_0 = 1.5, \Delta = 0.9 \mu m, \Delta = 0.031248$, and $x_0 = 60 \mu m$. For $\nu$ positive, the profile (9) describes a quadratic gain profile. An analytic solution to Eq. (2) for the index profile (9) can be written as

$$\psi(x, z) = \sum_{n=0}^{N} A_n \psi_n(x) \exp[-iv(1 - (1 + 2\beta_n^2/k^2)]z,$$

(10a)

$$\psi_n(x) = N_n \exp(-x^2/2\sigma^2) H_n(x/\sigma),$$

(10b)

$$\sigma = \left[ \frac{x_0(1 - 2\Delta)}{2k\Delta(1 - iv)^2} \right]^{1/2},$$

(10b)

where $N_n$ and $H_n$ are normalization constants and Hermite polynomials and $\beta_n$ are the paraxial mode propagation constants for the quadratic profile (9), namely,

$$\beta_n^2 = \frac{\Delta n_n^2(1 - iv)^2 \omega}{n_0} - \frac{n_1(1 - iv)(2\Delta)^{1/2}}{x_0} \left( n + 1 \right)^2.$$

(11)

The mode eigenfunctions (10b) and propagation constants (11) are accurate as long as modes are not too close to the cutoff.

Figure 1 shows the evolution of the beam as it propagates along the waveguide when the starting
beam is the finite delta function,

$$\psi(x,0) = \sum_{n=0}^{\infty} \phi(x_i) \phi(x),$$  \hspace{1cm} (12)

where the offset is \( x_1 = 20 \, \mu m \). Other parameters are \( \Delta x = 1.4 \, \mu m \), the number of transverse grid points is 128, \( \Delta z = 3.04 \, \mu m \), and the propagation distance \( z \) is expressed in units of the paraxial periodic focusing distance for the profile (9), namely, \( Z_f = 0.146 \, cm \). The parameter \( \nu \) was set to give the maximum gain length at the waveguide center the value 0.05 cm. Plotted in Fig. 1 are the beam patterns for the Helmholtz solution (solid curves) and the paraxial solution (dotted curves), both computed using the generalized Lanzcos orthogonalization scheme with \( N = 6 \), i.e., the equivalent of fifth order in a Taylor expansion sense.

The analytic and numerical Helmholtz solutions, which are superimposed, are indistinguishable. It was found that the maximum difference between the on-axis analytic and computed Helmholtz intensities was of the order of one part in \( 10^6 \).

Finally, in Fig. 2 the corresponding results obtained by using the same parameters but changing the sign of \( \nu \) to correspond to an absorption length of 0.05 cm are displayed. The accuracy for the absorption case was comparable with that for the gain case, and again the analytic and numerical curves for the Helmholtz equation solutions are indistinguishable. The discrepancy between paraxial and Helmholtz fields for large \( z \), owing to dephasing of the modes in the case of Helmholtz propagation, on the other hand, is even more pronounced than in the case of gain.

Comparisons between the previous formulation of Lanczos propagation and the generalized version described in this Letter, which space does not permit us to describe, indicate that the generalized method is capable of larger propagation steps and is in general more robust. In particular we have found that the generalized scheme correctly describes the coupling of energy between arms of a directional coupler, quite independent of the input field.

This research was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract W-7405-ENG-48.

References
8. By factoring out the carrier \( \exp(ikz) \) rather than \( \exp(-ikz) \), we make it easier to select the exponentially decaying modes that correspond to propagation angles greater than 90 deg (see Ref. 2).