ESD-TR-85-188

(2)

Project Report

LTP-10

J. Herrmann
L.C. Bradley

Numerical Calculation of Light Propagation

12 July 1971

Prepared for the Advanced Research Projects Agency
under Electronic Systems Division Contract F19628-70-C-0002 by

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

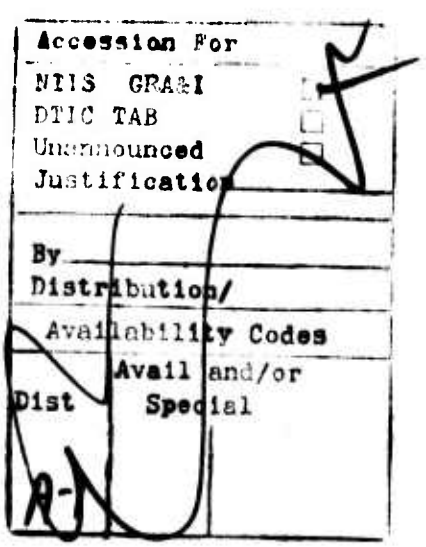
The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. This work was sponsored by the Advanced Research Projects Agency of the Department of Defense under Air Force Contract F19628-70-C-0002 (ARPA Order 600).

The ESD Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Thomas J. Alpert, Major, USAF
Chief, ESD Lincoln Laboratory Project Office



MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

NUMERICAL CALCULATION OF LIGHT PROPAGATION

J. HERRMANN

L.C. BRADLEY

Group 51

PROJECT REPORT LTP-10
(Laser Technology Program)

12 JULY 1971

LEXINGTON **MASSACHUSETTS**

ABSTRACT

An algorithm is presented for the numerical integration of the scalar wave equation in the Fresnel approximation in inhomogeneous media. The wave equation is reduced by a variable transform to an equation describing the deviation from a Gaussian reference beam. Large phase shifts are handled analytically in combination with the longitudinal discretization using the Crank-Nicolson scheme. This permits large step sizes. The discretization in the two transverse directions is performed with the Galerkin method, with spline functions as basis. An alternating-direction scheme is used in inverting the implicit finite-difference equations that result. An interaction of the longitudinal and transverse discretization is described. A method of handling a non-linear index of refraction is described, and the algorithm is applied to the thermal lens effect.

Although calculations of the propagation of a diffracting light beam have been made for over a century, almost all of them have been analytic approximations when exact procedures are not appropriate. Recently faced with a problem of nonlinear propagation to which an analytic solution seemed impossible without crippling assumptions, we have developed a finite difference algorithm for the calculation of light propagation which has proven capable of giving accurate results with reasonable economy of computation.

I. Reduction of the Wave Equation

As our starting point we take the scalar wave equation

$$\nabla^2 E - \frac{\varepsilon}{c^2} \frac{\partial^2 E}{\partial t^2} = 0 \ . \tag{1}$$

It has recently been shown[1] that the scalar equation gives correct results for electromagnetic propagation when polarization effects are not important. Here ε is in general a function of position, not necessarily real (i.e., absorption or stimulated emission may be present), and may also depend on the light intensity.

We define $\varepsilon = \varepsilon_o(1 + i\alpha/k) + \delta\varepsilon$, where $k^2 = \varepsilon_o \omega^2/c^2$ and ε_o is a constant. Here α may be a function of Z, though not of X and Y, and $\delta\varepsilon$ may be a complex function of all coordinates, and also of the field strength. Then we assume a monochromatic wave and remove the oscillations of highest spatial frequency by use

of the Ansatz

$$E = U(X,Y,Z)\ \exp[ikZ-i\omega t- \frac{1}{2}\int \alpha dZ] \ .$$

The beam is thus assumed to be propagating mainly in the Z - direction. If we now assume that the convergence (or divergence) of the beam is small, that $\alpha << k$ and that $\delta\varepsilon$ is not large enough to cause sharp deflections of a ray, we may neglect $\frac{\partial^2 U}{\partial Z^2}$ - the Fresnel approximation - α^2, $\frac{\partial \alpha}{\partial Z}$, and $\alpha\frac{\partial U}{\partial Z}$, and we are left with the equation

$$2ik\ \frac{\partial U}{\partial Z} + \frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} + k^2 \Delta\varepsilon\ U = 0 \ , \qquad (2)$$

where $\Delta\varepsilon = \delta\varepsilon/\varepsilon_o$.

For $\Delta\varepsilon = 0$ the solutions of this equation are well known[2]; they may be written

$$U_{mn} = \frac{1}{a}\ H_m(X/a)\ H_n(Y/a)\ \exp[-(X^2+Y^2)\,(\frac{1}{2a^2} + \frac{ik}{2R}) +$$

$$i\ (m + n + 1)\ \tan^{-1}\ Z/ka_o^2] \ , \qquad (3)$$

where

$$a^2 = a_o^2\ [1 + (Z/ka_o^2)^2] \ , \ R = Z\left[1 + \left(\frac{ka_o^2}{Z}\right)^2\right], \text{ and } H_m \text{ is a}$$

Hermite polynomial. The parameter a, which is the measure of the transverse scale, shrinks or expands as the beam converges to or diverges from the focus.

For numerical work with this problem, it is convenient to have a mesh that varies as the scale parameter a varies, since the cases of practical interest are those that do not deviate too grossly from the unperturbed case. Therefore we introduce the variables $x = X/a$, $y = Y/a$. Since the parameter $\Delta Z/(\Delta X)^2$ determines the stability and convergence of a numerical approximation to this parabolic equation, it is convenient to introduce a new axial variable such that this parameter automatically remains constant as the x and y meshes vary. This is accomplished by introducing $z = \tan^{-1} Z/ka_o^2$ and using a constant Δz. It may be noted that $\left(\Delta z/(\Delta x)^2\right)^{-1}$ is just the Fresnel number of a range Δz with respect to an aperture of radius Δx. Finally it is helpful to normalize out the rapid phase variations in the x-y plane, and the amplitude variations that come simply from the change of scale:

$$U = a^{-1} B \exp[ik(X^2+Y^2)/2R - i \tan^{-1} Z/ka_o^2]$$

$$= a_o^{-1} B \cos z \exp[i(x^2 + y^2) \tan z - iz] . \qquad (4)$$

With these changes in dependent and independent variables, the approximate wave equation becomes

$$2i \frac{\partial B}{\partial z} + \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2} + (2-x^2-y^2+k^2a^2\Delta\varepsilon)B$$

$$= 2i \frac{\partial B}{\partial z} - (L_x^+ L_x + L_y^+ L_y)B - G \quad B = 0 . \qquad (5)$$

Here the ladder operator $L_x = \frac{\partial}{\partial x} + x$, its adjoint $L_x^+ = -\frac{\partial}{\partial x} + x$, and $G = -k^2a^2\Delta\varepsilon$. It will be recognized that, aside from the term in $\Delta\varepsilon$, this is just the time-dependent Schrödinger equation for the two-dimensional simple harmonic oscillator, with the zero-point energy subtracted out, and the variable z corresponding to time. The complete range of z from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ corresponds to one half-period of the oscillator.

Another grouping of terms is possible and in some practical cases preferable. The choice $L_x' = \frac{\partial}{\partial x}$ and $G' = x^2 + y^2 - 2 - k^2a^2\Delta\varepsilon$ leads to numerical advantages for large values of $|x|$ or $|y|$. The general discussion that follows does not depend on which of these choices is made.

It is still possible to simplify Eq. (5) for numerical work, by recognizing that the term containing G can be removed analytically. For this purpose we introduce the new variable

$$D = B \exp\left(i/2\int^z G\, dz'\right) = B \exp(i\Gamma) , \qquad (6)$$

which is a unitary transformation if G is real. The lower limit of integration will be specified later. We shall also discuss later the problems that arise when G depends on B. The variable change Eq.(6) can be of considerable practical importance, reducing the number of longitudinal steps by as much as an order of magnitude.

Then the equation finally becomes

$$2i\,\frac{\partial D}{\partial t} - HD = 0 \qquad (7)$$

where the operator H is given by

$$H = \exp(i\Gamma)\ (L_x^+ L_x + L_y^+ L_y)\ \exp(-i\Gamma)\ . \qquad (8)$$

It is Eq. (7) that we approximate by finite differences.

Some workers have considered it desirable to choose another set of transverse variables so that the transverse boundaries are at a finite distance. We do not believe this is desirable, because the maximum step size in z is limited by the smallest physical mesh spacing, while the accuracy of discretization is limited by the largest physical mesh spacing in regions containing an appreciable fraction of the energy flux. We may however point out that if such a change is made, a simultaneous change should be made in the dependent variable, in such a way that the transverse differential operator remains in Sturm-Liouville form. Then the standard difference approximation will be a Hermitian operator, which is important for our numerical method as outlined below.

II. Numerical Method

The choice of a numerical method is of central importance, for practical reasons, and we shall therefore discuss this question in some detail. It is generally considered that complicated problems in partial differential equations are most readily approximated by finite difference equations, and this is the approach that we have used. We shall consider first the discretization in the longitudinal direction, then that in the transverse direction;

and finally we shall discuss methods of solving the resulting set of difference equations.

1. Longitudinal Discretization

We start by considering discrete coordinates in z: $z_{n+1} = z_n + \Delta z_n$, where the Δz_n need not all be the same. To advance in z we use the Crank-Nicolson algorithm:

$$D^{(n+1)} = \frac{1 - i\Delta z_n \bar{H}/4}{1 + i\Delta z_n \bar{H}/4} D^{(n)} \tag{9}$$

or

$$(1 + i\Delta z_n \bar{H}/4)\ D^{(n+1)} = (1 - i\Delta z_n \bar{H}/4)\ D^{(n)} \ . \tag{10}$$

Here D is still considered to be a continuous variable in the transverse directions. If H is an explicit function of z, some intermediate $\bar{H}$ must be used, as will be discussed later. The Crank-Nicolson approximation has a truncation error that is $O(\Delta z^3)$. Goldberg et al[3] have used this algorithm in the analogous case of the time-dependent Schrödinger equation. They point out that if H is Hermitian (which it will be in our case if G is real) the algorithm is unitary, i.e., the inner product (D,D) is conserved.* This corresponds to conservation of energy <u>flux</u> in our case, and to conservation of probability in the Schrödinger case.

*The inner product is defined as $(f,q) \int f^* q \, dx \, dy$.

The Schrödinger analogy suggests that other conservation laws are obeyed by the exact solution, and should be obeyed by a good approximate solution. In the Schrödinger case these are the laws of motion for energy and momentum; they correspond in our case to energy density and beam deflection. We distinguish between energy flux and energy density; when the index of refraction is not a constant, the integrated energy density is not a constant of the motion, but rather varies with the variation of light velocity in such a way that the flux, which is the product of velocity and energy density, is conserved. (In the Schrödinger case this corresponds to motion in a time-dependent potential.) To be exact, the law of motion for the integrated energy density is

$$\frac{d}{dz}\,(D,HD) = \left(D,\ \frac{\partial H}{\partial z}\,D\right) \tag{11}$$

The Crank-Nicolson algorithm leads to the following approximation to this equation of motion:

$$\left(D(z_{n+1}),\ H(z_{n+1})\ D(z_{n+1})\right) - \left(D(z_n),\ H(z_n)\ D(z_n)\right) =$$

$$\left(D(z_{n+1}),\ [H(z_{n+1}) - \bar{H}]\ D(z_{n+1})\right) + \left(D(z_n),\ [\bar{H} -$$

$$H(z_n)]\ D(z_n)\right).$$

This will be a good approximation if $\bar{H}$ is chosen to be some average of $H(z_n)$ and $H(z_{n+1})$. There are two obvious choices:

$\bar{H} = \frac{1}{2}\,[H(z_{n+1}) + H(z_n)]$ and $\bar{H} = H(z_n + \Delta z_n/2)$. Either is satisfactory, but we take the latter for reasons that will appear later.

Similarly, the law of motion for the integrated (transverse) momentum is

$$\frac{d}{dz}\, i\,(D,\nabla D) = \frac{1}{2}\,(D,\,[\nabla\bar{H} - \bar{H}\nabla]D).$$

The corresponding law, derived from equation (10) by multiplying both sides by $i\nabla(D^{(n)}+D^{(n+1)})$ and taking the real part, is

$$i[(D^{(n+1)},\,\nabla D^{(n+1)}) - (D^{(n)},\,\nabla D^{(n)})] = \frac{\Delta z}{2}\,(\bar{D},\,[\nabla\bar{H} - \bar{H}\nabla]\bar{D})$$

where $\bar{D} = \frac{1}{2}\,(D^{(n)} + D^{(n+1)})$. This is clearly a satisfactory form.

The usefulness of conservation laws in designing difference schemes has been emphasized by Morton[4], and some results on a scheme essentially the same as ours for equations of Schrödinger type have been derived by Kreiss.[5] It is known that conservation of energy flux is by itself sufficient to guarantee stability (and therefore convergence), in this case. This is because the conservation of the positive definite quantity (D,D) automatically limits the excursions of $|D|$. The importance of the law of motion of energy density which involves the positive definite quantity $(D,-\bar{H}D)$ has not previously been emphasized. It is clear though that a difference scheme that limits this quantity through a law of motion will automatically limit the excursions of $|\nabla D|$. This

may be particularly important where changes in D are largely due to changes in phase rather than amplitude. The momentum conservation law will be of smaller practical importance because the conserved quantity is not positive definite, so that cancellations can occur.

The fact that an algorithm obeys conservation laws is not a necessary condition for stability, as is evident from the fact that the method of Harmuth[6] does not do so; but it is a sufficient condition, and this is all that one ordinarily requires in practice. The Crank-Nicolson algorithm is stable for all values of Δz, while this is not true of the Harmuth algorithm. The limitation this implies on Δz is more apparent than real, though, since some limitation on Δz is required for accuracy in any case. Both schemes are $O(\Delta z^2)$; the coefficient of Δz^2 is not known a priori, but it appears experimentally that it is usually smaller for conservative schemes.[7] The principal advantage of the Harmuth scheme is that it is explicit, for the central difference approximation. Since this advantage is not retained when one uses the more accurate spline approximations for the transverse discretization discussed below, and since the Harmuth method requires twice as much storage, we have seen no reason to prefer it to the Crank-Nicolson scheme.

For higher accuracy in the longitudinal direction, one may use a generalized Crank-Nicolson scheme proposed by Nassif[8] and

applied by him to the diffusion equation. He points out that the usual Crank-Nicolson scheme is simply the 1,1 Padé matrix approximation of $\exp(-i\, z_n \bar{H}/2)$, and that the higher approximations may be found by using higher order Padé approximants.[9] He further observes that the calculation becomes computationally practicable if the approximants are factored. For our problem this approach would be entirely feasible as long as diagonal (e.g., 2,2 or 3,3) approximants are used, for in this case all the conservation laws would be satisfied. The approach would perhaps even be more natural in our problem than in the diffusion problem, for the factorization introduces complex numbers, which are not originally present in the latter. We have, however, made no attempt to apply Nassif's refinement to our problem.

We may now specify the lower limit to the integral in Eq. (6). This limit we take to be at the half-way point $z_n + \Delta z_n/2$, so that in effect we make the variable change Eq. (6) anew at every step in z. This leads to $\Gamma = 0$ at the halfway point, and thus from Eq. (8) $\bar{H} = L_x^+ L_x + L_y^+ L_y$. It is evident that with this choice $\bar{H}$ is always Hermitian, and the algorithm will always be stable even if G is not real, i.e., if absorption or stimulated emission are present.

We may then rewrite Eq. (10) in terms of the variable B, which is the one used in actual computation:

$$(1+i\Delta z_n \bar{H}/4)\, \exp(i\Gamma_{n+1})\, B^{(n+1)} = (1-i\Delta z_n \bar{H}/4)\, \exp(i\Gamma_n)\, B^{(n)} \qquad (12)$$

$$\text{where } \Gamma_n = \frac{1}{2}\int_{z_n+\Delta z_n/2}^{z_n} G\,dz\ , \qquad \Gamma_{n+1} = \frac{1}{2}\int_{z_n+\Delta z_n/2}^{z_{n+1}} G\,dz. \tag{12a}$$

2. Transverse Discretization

So far we have considered B and D to be continuous functions of x and y. We now proceed to impose a mesh in the x-y plane, with mesh lines at $x = x_i$, $y = y_j$. But instead of making the usual central-difference approximation for the derivatives at the mesh points, we approximate D by spline functions:

$$D \approx \tilde{D} = \sum_{i,j} D_{ij}\, w_i(x) w_j(y)\ . \tag{13}$$

The spline basis functions $w_i(x)$ are defined by

$$w_i(x) = 0 \quad \text{if} \quad x \le x_{i-1} \quad \text{or} \quad x \ge x_{i+1}$$

$$w_i(x_i) = 1\ . \tag{14}$$

The behavior in the intervals $x_{i-1} \le x \le x_i$ and $x_i \le x \le x_{i+1}$ may be chosen in any convenient way, but only two choices are of interest to us here. The simplest choice specifies that

$$\frac{d^2}{dx^2} w_i(x) = 0 \quad \text{except at mesh points.} \tag{15}$$

This leads to the so-called roof function or chapeau function, the linear spline basis function, which corresponds in one dimension to simple straight-line interpolation between mesh points. It is suitable when all regions of the plane are to be treated equally.

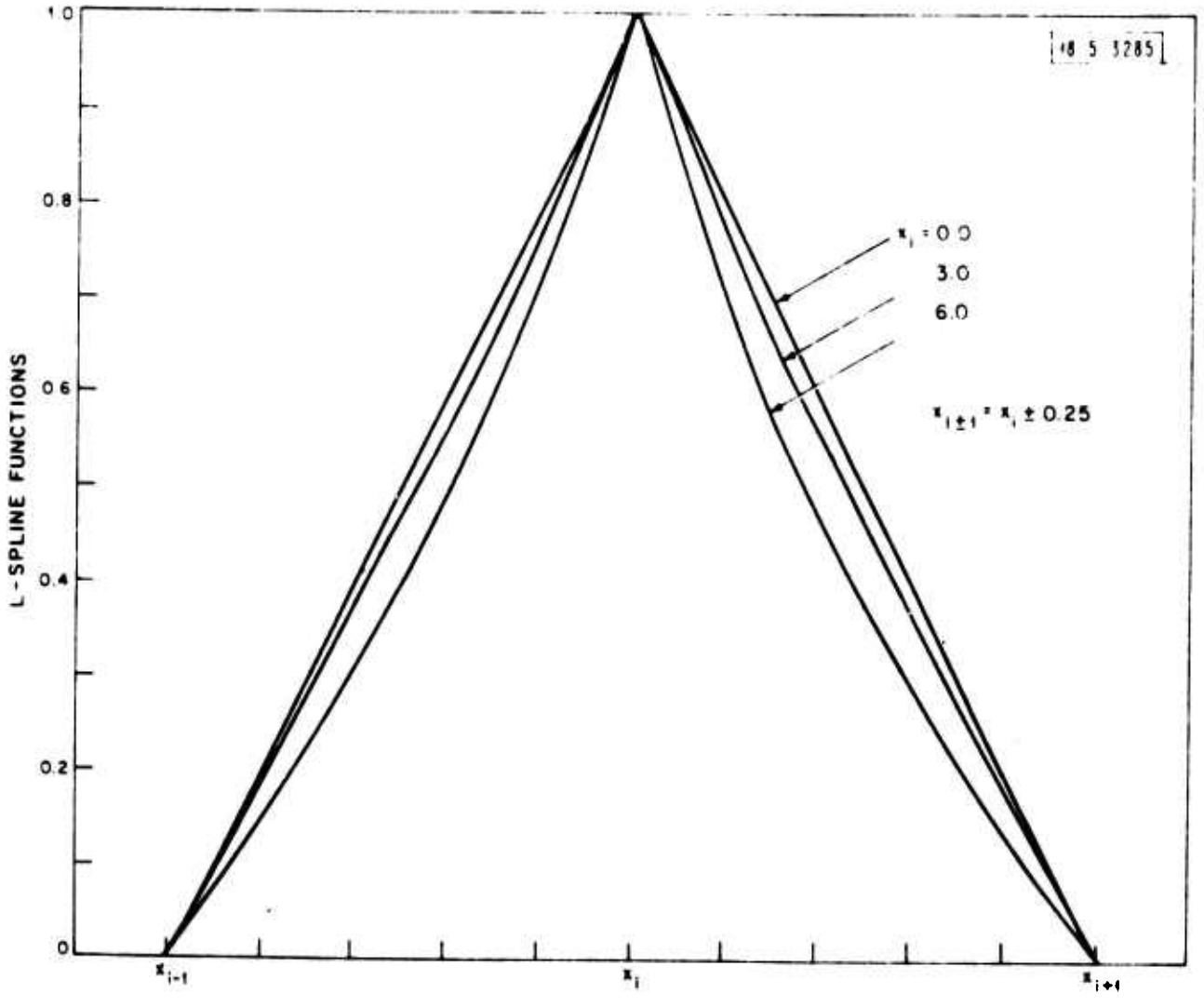


Fig. 1. L-splines for the operator $(\partial^2/\partial x^2 + 1 - x^2)$, centered at various x_i.

But our coordinate transformation has led to an equation which, to speak in Schrödinger language, has a central potential. This leads to fields that decrease rapidly at large distances, as is evident from the eigenfunctions of Eq. (3); for fields that do not do so the Fresnel approximation is not valid. In this situation the so-called L-spline[10,11] basis functions are more appropriate:

$$L_x^+ L_x \, w_i(x) = 0 \quad \text{for } x \text{ not a mesh point.} \tag{16}$$

These functions are very similar to the chapeau function for small $|x|$, but have increasing curvature for increasing $|x|$, and thus can approximate a rapidly decreasing function more accurately. In fact they can reproduce exactly a Gaussian centered at the origin. Examples of these L-spline basis functions are shown in Fig. 1. For a finite number of mesh points the end segments (with either choice) are treated somewhat differently:

$$\begin{aligned} L_x w_1(x) &= 0 \quad \text{for} \quad -\infty \le x \le x_1 \\ L_x w_N(x) &= 0 \quad \text{for} \quad x_N \le x \le \infty \end{aligned} \tag{17}$$

These conditions lead to the form

$$\begin{aligned} w_1(x) &= \exp \tfrac{1}{2}(x_1^2 - x^2) \quad \text{for} \quad -\infty \le x \le x_1 \\ w_N(x) &= \exp \tfrac{1}{2}(x_N^2 - x^2) \quad \text{for} \quad x_N \le x \le \infty \end{aligned} \tag{18}$$

which automatically satisfy the boundary conditions at infinity.

At this point we should say a little more about spline functions, since they are not well known outside the applied mathematics community. Polynomial spline functions were introduced by Schoenberg[12], and have been extensively used for interpolation and approximation, with cubic splines being favored. Although a number of books have appeared dealing with spline functions and spline approximation[13,14,15], these have been written in the language of functional analysis and are not addressed to the general technical reader. Furthermore, except for the articles of Jerome and Varga in Ref. 14 and of Schultz in Ref. 15, they hardly touch on the use of splines in the solution of partial differential equations. The most important results for our purposes are that polynomial spline approximations using polynomials of degree (2m-1) will converge to a given function with an error (in some norm) that is $O(h^{2m})$, where h is the mesh spacing, if the function satisfies certain continuity conditions; and that this form of approximation is particularly suited to approximating functions that are rapidly changing. Thus these functions can be used as basis functions in a Ritz or Galerkin calculation, since the requirement on such basis functions is essentially that they be capable of forming an approximating sequence that converges to the desired function.

We apply the Galerkin method to Eq. (10):

$$\left(w_i(x)w_j(y),\ [1+i\Delta z_n\bar{H}/4]\ \tilde{D}^{(n+1)}\right) = \left(w_i(x)w_j(y),\ [1-i\Delta z_n\bar{H}/4]\tilde{D}^{(n)}\right) \tag{19}$$

for all i,j. This is equivalent to a Ritz variational procedure if H is self-adjoint, but is also applicable to more general cases.† It may be rewritten in matrix form in the following way. Define the matrices C_x, H_x as having the elements

$$(C_x)_{ik} = (w_i, w_k), \quad (H_x)_{ik} = (w_i, L_x^+ L_x w_k) = (L_x w_i, L_x w_x).$$

Then if the vector $\hat{D}$ is defined as having the components D_{ij}, equation (19) becomes, in matrix form,

$$[C_x \otimes C_y + \frac{i\Delta z_n}{4} (C_x \otimes H_y + H_x \otimes C_y)] \hat{D}^{(n+1)}$$

$$= [C_x \otimes C_y - \frac{i\Delta z_n}{4} (C_x \otimes H_y + H_x \otimes C_y)] \hat{D}^{(n)} , \qquad (20)$$

where the symbol $\otimes$ denotes the direct product.

Provided that H is Hermitian the conservation laws for energy flux and energy density are preserved by Eq. (19), in the matrix form

$$\hat{D}^{+(n+1)} (C_x \otimes C_y) \hat{D}^{(n+1)} = \hat{D}^{+(n)} (C_x \otimes C_y) \hat{D}^{(n)}$$

and

$$\hat{D}^{+(n+1)} \hat{H}^{(n+1)} \hat{D}^{(n+1)} - \hat{D}^{+(n)} \hat{H}^{(n)} \hat{D}^{(n)} =$$

$$\hat{D}^{+(n+1)} [\hat{H}^{(n+1)} - \bar{\hat{H}}] \hat{D}^{(n+1)} + \hat{D}^{+(n)} [\bar{\hat{H}} - \hat{H}^{(n)}] \hat{D}^{(n)}$$

---

†Our procedure is essentially that of reference 11, and has much in common with the finite element method[16] and the work of Rose.[17] The use of spline functions in conjunction with the Galerkin method has the great advantage for numerical work that the matrices to be inverted are sparse.

with the matrix $H^{(n)}$ $C_x \otimes H_y^{(n)} + H_x^{(n)} \otimes C_y$, and so on.

The law of motion for momentum is not similarly preserved in general. Defining an x-momentum matrix

$$(P_x)_{ik} = i\int w_i \frac{d}{dx} w_k \, dx = (P_x)^*_{ki} ,$$

one can derive an equation of motion that involves the symmetrized matrix $P_{xs} = \frac{1}{2}(P_x C_x + C_x P_x) \otimes C_y$ and the commutator $P_{xa} = \frac{1}{2}(P_x C_x - C_x P_x) \otimes C_y$:

$$\hat{D}^{+(n+1)} P_{xs} \hat{D}^{(n+1)} - \hat{D}^{+(n)} P_{xs} D^{(n)}$$

$$\frac{i\Delta z}{2} \bar{D}^{+} (P_x \Omega_x - \Omega_x P_x) \otimes C_y \bar{D} + \hat{D}^{+(n+1)} P_{xa} D^{(n)} - \hat{D}^{+(n)} P_{xa} \hat{D}^{(n+1)} ,$$

where $\bar{D} = \frac{1}{2}(\hat{D}^{(n)} + \hat{D}^{(n+1)})$. The terms containing P_{xa}, which represent a deviation from the desired form, vanish if one uses linear splines for the w_i. They do not vanish if one uses L-splines, but they are small. Because the momentum law does not deal with the conservation of a positive definite quantity, we do not believe this defect is of practical importance.

3. Discretization Interactions

The effect of using the approach outlined above instead of simply substituting the central-difference approximation for the second derivatives may be illustrated by considering a simpler example. The equation

$$2i \frac{\partial D}{\partial z} + \frac{\partial^2 D}{\partial x^2} = 0$$

would be approximated, using the central-difference scheme, by

$$(I - \frac{i\eta}{2} \delta^2)\ \hat{D}^{(n+1)} = (I + \frac{i\eta}{2} \delta^2)\ \hat{D}^{(n)} \quad . \tag{21}$$

Here the matrix operator δ^2 is defined by $(\delta^2 D)_i = D_{i+1} - 2D_i + D_{i-1}$, I is the unit matrix and $\eta = \Delta z / 2h^2$. With the Galerkin method and linear splines, the analogous approximation is

$$(C - \frac{i\eta}{2} \delta^2)\ \hat{D}^{(n+1)} = (C + \frac{i\eta}{2} \delta^2)\ \hat{D}^{(n)} \quad , \tag{22}$$

where the operator $C = I + \frac{1}{6} \delta^2$. This comparatively small change makes a considerable improvement in the accuracy of the integration without any addition to the computation time. We remark in passing that it is not equivalent to fourth-order accuracy in the x-direction, which would be achieved by the choice $C' = I + \frac{1}{12} \delta^2$. This latter choice would presumably converge more rapidly for small mesh sizes, but is less accurate for larger mesh sizes. An analogy to this is the difference between Taylor-series approximation and Chebyshev-economized approximation; for a given number of terms the former is more accurate near the origin, but the latter has a smaller maximum error over a wide range.

The advantage of Eq. (22) over Eq. (21) may be illustrated by considering the propagation of the function e^{ikx}, which is an eigenfunction for both the differential and the various finite-difference equations. The exact solution to the differential

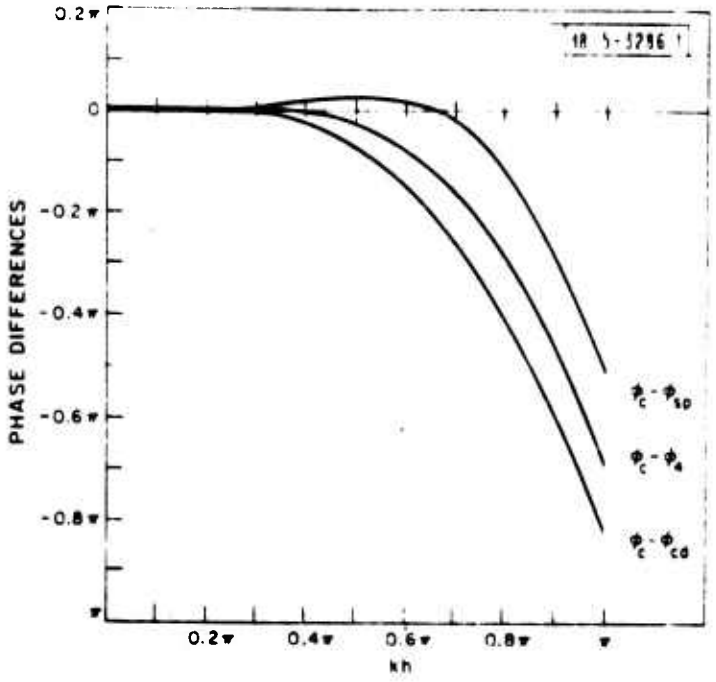


Fig. 2. Phase deviations of three differencing schemes for $\eta \equiv \Delta z/2h^2 = 0.4$.

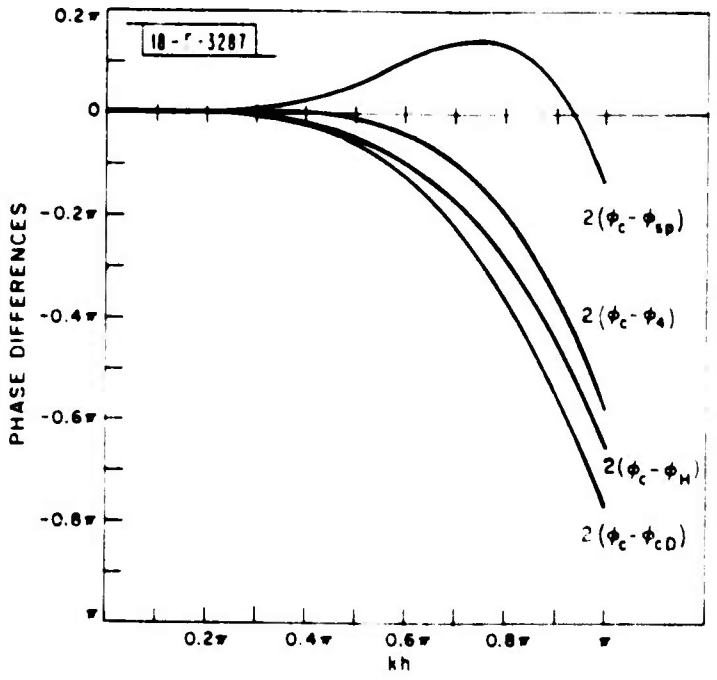


Fig. 3. Phase deviations of four differencing schemes for $\eta = 0.2$.

equation gives a phase change $\phi_c = -\eta(kh)^2$ for propagation through a distance Δz. For the finite difference schemes k assumes only discrete values, with a maximum $kh = \pi$, corresponding to the Nyquist frequency.

The corresponding phase change ϕ_{cd} for equation (21) can be found by observing that

$$\delta^2 \exp(ikx_i) = \lambda \exp(ikx_i) \quad \text{with} \quad \lambda = -4 \sin^2 \frac{kh}{2} ,$$

so that

$$\exp(i\phi_{cd}) = \frac{1 + i\eta\lambda/2}{1 - i\eta\lambda/2} .$$

This leads to the result

$$\phi_{cd} = 2 \tan^{-1} (\eta\lambda/2) .$$

A similar calculation using equation (22) yields

$$\phi_{sp} = 2 \tan^{-1}\left(\frac{\eta\lambda/2}{1 + \lambda/6}\right)$$

and the fourth-order choice C' leads to

$$\phi_4 = 2 \tan^{-1}\left(\frac{\eta\lambda/2}{1 + \lambda/12}\right) .$$

The differences $\phi_c - \phi_{cd}$, $\phi_c - \phi_4$, and $\phi_c - \phi_{sp}$ are plotted in Fig. 2 for $\eta = 0.4$. In order to show the effect of a change of η, in Fig. 3 we plot $2(\phi_c - \phi_{cd})$, $2(\phi_c - \phi_4)$, and $2(\phi_c - \phi_{sp})$ for $\eta = 0.2$. These curves thus correspond to the same advance in z as the curves of Fig. 2. In both cases it is clear that ϕ_{sp} is a better approximation than ϕ_{cd} or ϕ_4 at high spatial frequencies. It is also clear that for intermediate frequencies $\eta = 0.4$ is a better choice than $\eta = 0.2$, while the opposite is true at high

frequencies.* In either case, the implication is that Eq. (22) can be used with considerably larger values of h than Eq. (21). Since the computing time is approximately proportional to h^{-4}, this is an important consideration.

We may also define the corresponding quantity for the Harmuth method[6]

$$\phi_H = \pm \sin^{-1}(n\lambda).$$

In Fig. 3 we also plot $2(\phi_C - \phi_H)$. The Harmuth method is unstable for $n > 0.25$ and the phase shift is therefore not included in Fig. 2.

4. Practical Considerations

Direct inversion of the matrix Eq. (20) would be very time-consuming. Instead we use the alternating-direction approximation[18,19] in the following form:

$$(C_x + \frac{i\Delta z}{4} H_x) \otimes (C_y + \frac{i\Delta z}{4} H_y)\, \hat{D}^{(n+1)} = (C_y - \frac{i\Delta z}{4} H_y) \otimes (C_x - \frac{i\Delta z}{4} H_x)\, \hat{D}^{(n)} \tag{23}$$

Since all the matrices in this equation are tridiagonal, and furthermore are diagonally dominant, the standard algorithm[20] for Gaussian elimination is available and is numerically stable.

---

*Although the interaction between longitudinal and transverse mesh sizes is considered in the usual stability analyses of finite-difference schemes, we are not aware of any general discussion in the literature of the influence of this interaction on the relative "fidelity" of such schemes.

Although the numerical scheme of Eq. (23) is stable for all η a small value of $\eta(\lesssim 0.5)$ is necessary in order to insure a reasonably accurate calculation of phase change, as we have seen in Section II.3. This is equivalent to the requirement that Δz be less than one quarter of the Rayleigh range of a disc of radius h.

The use of the alternating-direction approximation in the form of Eq. (23) means that somewhat different conservation laws are obeyed, but in practical cases it turns out that the differences are small.

When $\Delta\varepsilon$ is a given function of space the computation of Γ for each step presents no difficulties. But when $\Delta\varepsilon$ depends on B the problem is not so easy. Fortunately in most problems concerned with monochromatic radiation (as opposed to harmonic generation, for example), the dependence is on $|B|^2$. Since the variables D and B differ only by a phase change, a knowledge of $D^{(n+1)}$ is sufficient to permit the calculation of $\Delta\varepsilon^{(n+1)}$. we make the approximations

In the integrations leading to an evaluation of the Γ's, Eq. (12a), the z-variation of G is assumed to be only due to the variation of a(z), which in nonlinear cases also appears in $\Delta\varepsilon$.

Although our approximations for $\Gamma^{(n)}$ and $\Gamma^{(n+1)}$ have truncation errors that are $O(\Delta z^2)$, the symmetry of Eq. (12) leads to a truncation error in the overall process that is only $O(\Delta z^3)$.

We have used both linear splines and L-splines as basis functions, both with the end conditions of Eq. (18). Although in principle the L-spline functions are more appropriate, in practice we have found little difference in the results. In cases where $\Delta\varepsilon = 0$, so that one can compare with analytic expressions, the L-splines are more accurate, but we have considered that the differences are not usually enough to warrant the extra effort.

III. Application to the Thermal Lens Effect

The method of integration described in the previous sections has been used to calculate the propagation of a beam in an absorbing medium in a transverse wind.

For this physical situation, which is discussed in some detail by Wallace and Camac,[21] the dielectric constant can be written

$$\varepsilon/\varepsilon_o = 1 + i\alpha/k + \delta\varepsilon/\varepsilon_o ,$$

where the absorption of energy leads to

$$\delta\varepsilon = \frac{\partial\varepsilon}{\partial T}\,\frac{\alpha}{\rho c_p}\,\frac{1}{v}\int_{-\infty}^{x} I(x,y)\,dx.$$

Here α is the absorption coefficient, T the temperature, v the velocity of the wind in the x-direction, and I the flux density.

An example of the calculations using the experimental conditions of Kleiman and O'Neil[22] is presented in Figs. 4 and 5, which show intensity contours calculated for the same physical parameters, and with the same mesh, by the central-difference approximation and the Galerkin method, with spline functions as basis. That the spline calculation (Fig. 5) is the more accurate is shown by a comparison with the results when the mesh is refined. The results of the central-difference calculation then change in such a way that they approach more closely the spline calculation. The results of the spline calculation with a finer mesh are shown in Fig. 6. Although both methods presumably converge quadratically to the same result, it is clear that the use of splines gives a more accurate calculation for a given mesh spacing. Since the computing time goes approximately as h^{-4}, this gain is of considerable practical importance.

Other workers[23,24,25,26] have used a variety of schemes to calculate the thermal lens effect by integrating the wave equation. On the basis of preliminary reports we believe our method is far more economical for a given accuracy.

IV. Conclusions

As the example shows, our method works quite successfully. We have applied it to many cases of interest. But it should be pointed out that if the wavefront is severely distorted from the reference wavefront, the resulting phase difference, even if it

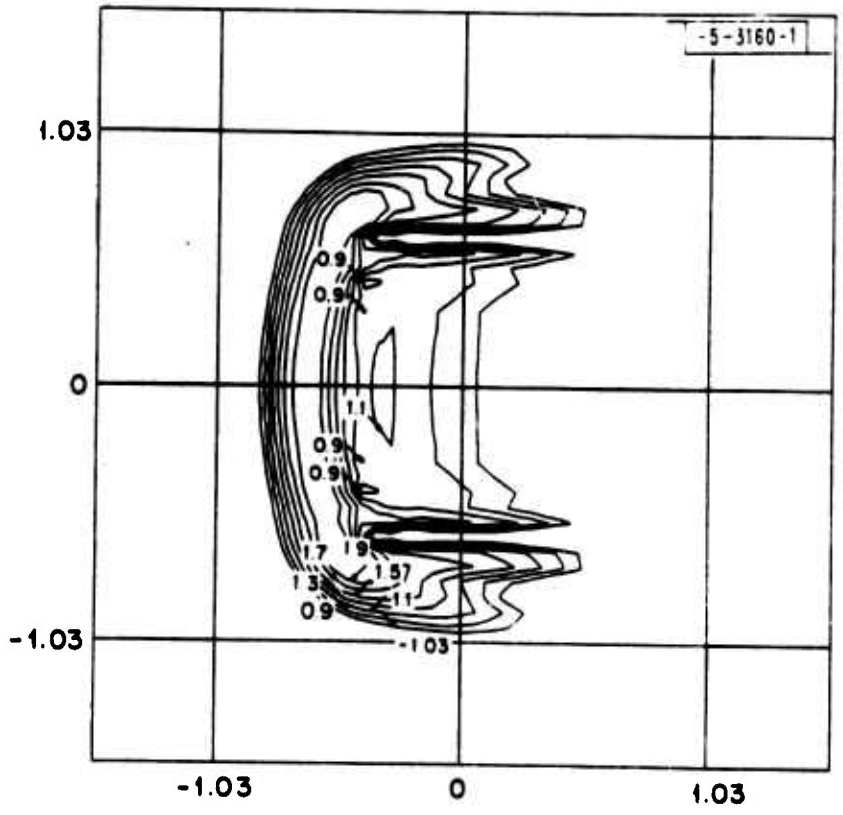


Fig. 4. Contour plot of intensity after passing through thermally blooming medium with an absorption path of 3 m. Contour spacing 2 dB. Beam power 70 W, diameter 7 mm, absorption coefficient 6×10^{-4} cm^{-1}, wavelength 10.6 μm. Transverse mesh spacing h = 0.25. Central difference approximation. Maximum intensity is 114.4 W/cm^2. Number of z-steps is 16.

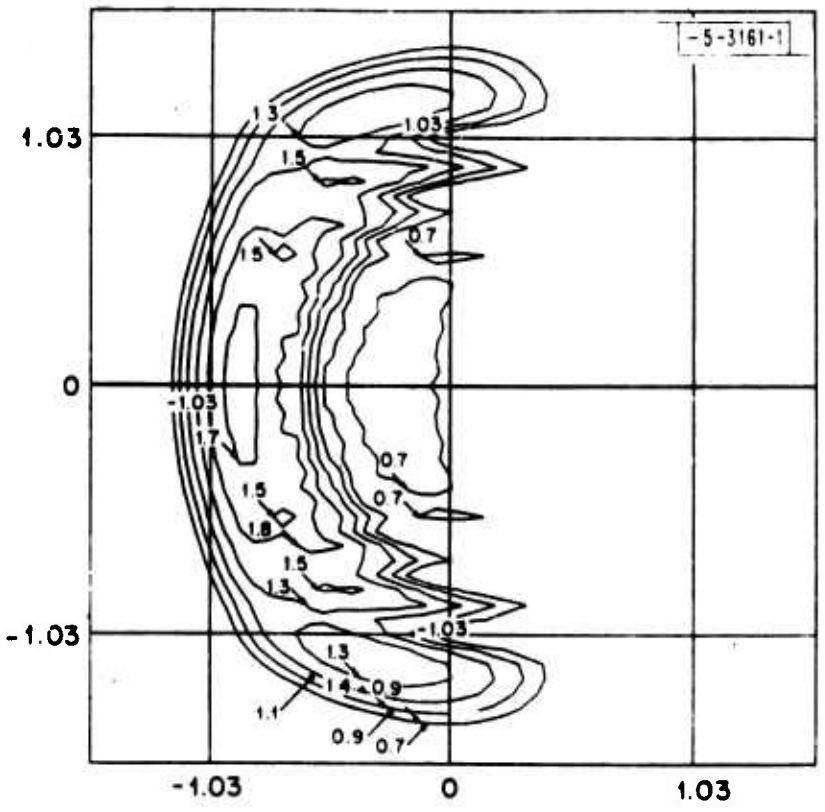


Fig. 5. Same as Fig. 4, spline approximation. Maximum intensity 62.0 W/cm^2.

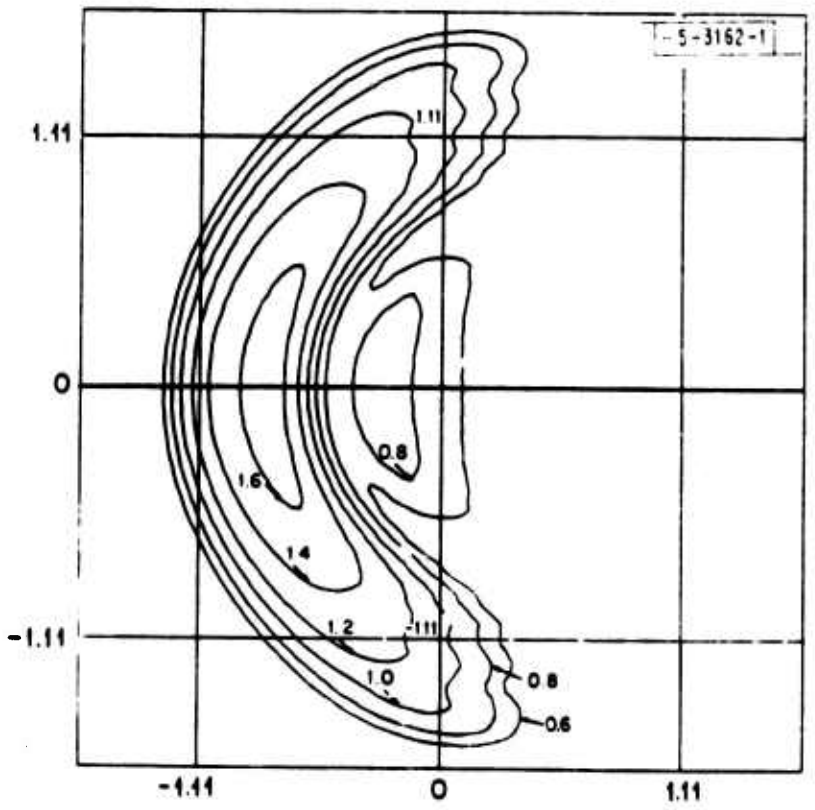


Fig. 6. Same as Fig. 5, mesh spacing h = 0.15, spline approximation. Maximum intensity 48.6 W/cm^2. Number of z-steps is 28. Required 6 minutes of computing time on an IBM 360-67.

be a smooth function of the transverse coordinates, will lead to rapid fluctuations of B in the x-y plane, and therefore will require a refined mesh. This requirement is particularly severe at short wavelengths. For such a problem it is likely that (real) amplitude and phase (instead of the complex amplitude) are the most suitable variables, but it is not so easy to find a stable numerical method for the solution of the resulting nonlinear equations. For less severe cases the present method seems good enough, though some economy in time might be achieved by combining Nassif's refinement of the Padé approximant with third-order Hermite L-splines. The experience of Ref. 11 suggests, though, that this economy would be achieved only in cases where high accuracy was required.

References

1. A. S. Marathay and G. B. Parrent, Jr., JOSA 60, 243 (1970).

2. H. Kogelnik and T. Li, Proc. IEEE 54, 1312 (1966).

3. A. Goldberg, H. M. Schey, and J. L. Schwartz, Am. J. Phys. 35, 177 (1967).

4. K. W. Morton, in Numerical Solution of Field Problems in Continuum Physics, G. Birckhoff and R. S. Varga, eds, American Mathematical Society, Providence, R.I., (1970).

5. H. O. Kreiss, Num. Math. 5, 24 (1963).

6. H. F. Harmuth, J. Math. Phys. 36, 269 (1957).

7. See, e.g., W. F. Ames, Non-linear Partial Differential Equations in Engineering (Academic Press, New York, 1962), p. 453.

8. N. Nassif, private communication.

9. R. S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, (1962), p. 266.

10. M. H. Schultz and R. S. Varga, Numer. Math. 10, 345 (1969).

11. H. S. Price and R. S. Varga, in Numerical Solution of Field Problems in Physics, G. Birckhoff and R. S. Varga, eds., American Mathematical Society, Providence, R.I., 1970.

12. I. J. Schoenberg, Quart. J. Appl. Math. 4, 45, 112 (1946).

13. J. H. Ahlberg, E. N. Nilson, J. L. Walsh, The Theory of Splines and Their Applications (Academic Press, New York, 1967).

14. T. N. E. Greville, ed. Theory and Applications of Spline Functions (Academic Press, New York, 1969).

15. I. J. Schoenberg, ed. Approximations with Emphasis on Spline Functions (Academic Press, New York, 1969).

16. O. C. Zienkiewicz, The Finite Element Method in Structural and Continuum Mechanics (McGraw-Hill, London and New York, 1967).

17. M. E. Rose, Math. Comp. 18, 179 (1964).

18. Ref. 9, p. 209, 273.

19. J. Douglas, Jr., and T. Dupont, "Alternating-Direction Galerkin Methods on Rectangles," to appear.

20. Ref. 9, p. 194.

21. J. Wallace and M. Camac, J. Opt. Soc. 60, 1587 (1970).

22. H. Kleiman and R. W. O'Neil, J. Opt. Soc. 61, 12 (1971).

23. J. Wallace, private communication.

24. A. H. Aitken, J. N. Hayes, and P. B. Ulrich, J. Opt. Soc. Am. 61, 674 (1971), IEEE J. Quantum Electronics, QE-7, 275 (1971), and private communication.

25. H. Breaux, private communication.

26. R. Butts, B. Hogge and W. Trebilcock, private communication.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-85-188	2. GOVT ACCESSION NO. AD-A954864	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Numerical Calculation of Light Propagation		5. TYPE OF REPORT & PERIOD COVERED Project Report 6. PERFORMING ORG. REPORT NUMBER Project Report LTP-10
7. AUTHOR*(s)* Jan Herrmann and Lee C. Bradley		8. CONTRACT OR GRANT NUMBER*(s)* F19628-70-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, MIT P.O. Box 73 Lexington, MA 02173-0073		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element No.62301E ARPA Order No. 600
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, VA 22209		12. REPORT DATE 12 July 1971 13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)* Electronic Systems Division Hanscom AFB, MA 01731		15. SECURITY CLASS. *(of this Report)* Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimted.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

laser propagation
atmospheric propagation
wave optics propagation

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

An algorithm is presented for the numerical integration of the scalar wave equation in the Fresnel approximation in inhomogeneous media. The wave equation is reduced by a variable transform to an equation describing the deviation from a Gaussian reference beam. Large phase shifts are handled analytically in combination with the longitudinal discretization using the Crank-Nicolson scheme. This permits large step sizes. The discretization in the two transverse directions is performed with the Galerkin method, with spline functions as basis. An alternating-direction scheme is used in inverting the implicit finite-difference equations that result. An interaction of the longitudinal and transverse discretization is described. A method of handling a non-linear index of refraction is described, and the algorithm is applied to the thermal lens effect.