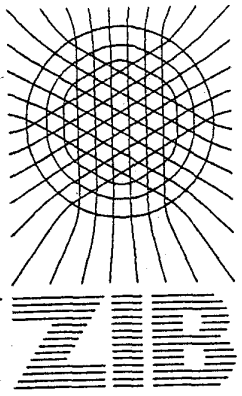
Konrad-Zuse-Zentrum für Informationstechnik Berlin



F. SCHMIDT

An Adaptive Approach to the Numerical Solution of Fresnel's Wave Equation

Preprint SC 91-12 (December 1991)

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Heilbronner Str. 10
1000 Berlin 31
Verantwortlich: Dr. Klaus André
Umschlagsatz und Druck: Rabe KG Buch-und Offsetdruck Berlin

ISSN 0933-7911

An Adaptive Approach to the Numerical Solution of Fresnel' s Wave Equation

F. SCHMIDT

Konrad–Zuse–Zentrum für Informationstechnik Berlin, Heilbronner Strasse 10, D-1000 Berlin 31, Federal Republic of Germany

ABSTRACT

An adaptive approach to the numerical solution of the wave propagation in integrated optics devices with 1D cross sections is described. First, Fresnel's approximation of the exact wave equation resulting from Maxwell's equations is considered. A criterion to estimate the validity of this approximation is derived. Fresnel's wave equation being formally equivalent to Schroedinger's equation uniquely defines an initial-boundary-value problem, which is solved numerically by a stepwise calculation of the propagating field.
Discretization in longitudinal direction first with stepsize control leads to a stationary subproblem for the transversal field distribution, which is then handled by an adaptive finite element method. Thus full adaptivity of the algorithm is realized.
The numerical examples are concentrated on taper structures playing an essential role in integrated optics devices for telecommunication systems.

Contents

1. INTRODUCTION

The interest in the numerical simulation of the wave propagation in integrated optics devices for optical telecommunication systems is constantly growing and much work is being invested to the development of fast and reliable algorithms. Starting with the basic publication of FEIT and FLECK [6] a large variety of methods have been published until now (see for example [15], [7], and [11]). Because we are interested in the handling of complicated spatial refractive index geometries, and the progress in the technology forces the trend to more sophisticated optical components, the algorithm proposed by KOCH [10] using a finite element method for the transversal field description was the starting point to develop a more efficient algorithm.

All the methods mentioned above are based on a discretization, which is fixed in advance, but some of the complicated components, for example the taper structures treated in this paper, can no longer be simulated efficiently by using an a-priori fixed discretization.

A successful handling of such devices requires an essential reduction of the computational amount of work, which becomes most evidently in the case of three space dimensions. An attractive concept of such an amount of work reduction is *adaptivity*, i. e., the automatic nonuniform discretization of every transverse cross section and the use of a related step size control in the longitudinal direction. The objective of this paper is to test the idea of adaptivity in the field of wave propagation in integrated optics devices. The algorithmic proposals concerning the adaptive formulation, whose application to optical problems are presented, are mainly due to the work of BORNEMANN [2], [3] and they are theoretically grounded there. Because this paper is devoted to the principal formulation of wave propagation using adaptive methods, it is, for the time being, restricted to problems with 1D cross sections. Of course, this restriction is understood as a first step.

The *physical* problem we want to solve can be described as follows. Suppose two infinite, parallel planes bounding the part of our component, we are intersted in. A coherent light source emits a beam, which travels through the first plane, the input plane and then through the component until it passes the second plane, the output plane. We are looking for the field distribution between both planes. In the most general case we have to solve the time dependent Maxwell's equations until a stationary state is reached. But if we would restrict ourselves to perfect reflection free structures, we will be able to formulate our problem as an initial-boundary-value problem, because every field distribution at a given plane parallel to the input plane defines

uniquely the field distribution behind this plane. Now, for practical purposes it should be enough, if the reflections within our component are sufficiently small. To formulate such an initial-boundary-value problem we will use the slowly varying amplitude approximation leading to Fresnel's wave equation, which is a standard procedure in optics (see e. g. [10]). But we will take some care in looking for an approximation, which is well suited for the adaptive numerical formulation. It will turn out, that the adaptive method reacts very sensitive to the 'inherent difficulty' of the problem and that therefore a very close interdependence between the modeling of the problem, the formulation of the algorithm, and the numerical efficiency can be observed.

2. The Paraxial Wave Equations

2.1. The Vector Wave Equations

It is assumed that the dielectric media is isotropic, nonmagnetic, and sourceless and that the permittivity ϵ depends only on the position. The usual time dependence $\exp(i\omega t)$ is understood, where ω denotes the optical angular frequency. Maxwell's equations are given by

$$\nabla \times \mathbf{H} = i\epsilon_0\epsilon_r\omega\mathbf{E} \tag{2.1}$$

$$\nabla \times \mathbf{E} = -i\mu_0\omega\mathbf{H} \tag{2.2}$$

$$\nabla \cdot \mathbf{H} = 0 \tag{2.3}$$

$$\nabla \cdot (\epsilon_r\mathbf{E}) = 0 \quad . \tag{2.4}$$

($\mathbf{H}$: magnetic field, $\mathbf{E}$: electric field, ϵ_0: permittivity in vacuum, ϵ_r: relative permittivity, μ_0: permeability in vacuum)

Using these equations, we get the vector wave equation for the electric and the magnetic field

$$\nabla^2\mathbf{E} + \omega^2\epsilon_0\epsilon_r\mu_0\mathbf{E} + \nabla\left(\frac{\nabla\epsilon_r}{\epsilon_r}\cdot\mathbf{E}\right) = 0 \quad , \tag{2.5}$$

$$\nabla^2\mathbf{H} + \omega^2\epsilon_0\epsilon_r\mu_0\mathbf{H} + \frac{\nabla\epsilon_r}{\epsilon_r}\times(\nabla\times\mathbf{H}) = 0 \quad . \tag{2.6}$$

We define the Poynting vector as usual to

$$\mathbf{S} = \mathbf{E}\times\mathbf{H}^* \quad . \tag{2.7}$$

where the '$*$' indicates the complex conjugate vector. From the equations (2.1) and (2.2) follows

$$\begin{aligned} \nabla\cdot\mathbf{S} &= \nabla\cdot(\mathbf{E}\times\mathbf{H}^*) \\ &= i\epsilon_0\epsilon_r\omega\mathbf{E}\mathbf{E}^* - i\mu_0\omega\mathbf{H}\mathbf{H}^* \quad . \end{aligned} \tag{2.8}$$

The theorem of Poynting states that the real part of $\mathbf{S}$ is the average power flux density [12].
We use a x, y, z-Cartesian coordinate system, where every plane $z = const$ defines a cross section of our integrated optical component with the transverse coordinates x and y, and the wave should travel almost parallel to the z-axis.

The integration of $\nabla \cdot \mathbf{S}$ over a cross section and the application of the divergence theorem in two dimensions gives

$$\int_\Omega \nabla \cdot \mathbf{S}\, d\Omega = \frac{\partial}{\partial z}\int_\Omega S_z\, d\Omega + \int_\Gamma \mathbf{S}_t \mathbf{n}\, d\Gamma \quad , \tag{2.9}$$

where Ω indicates an area integration in the x, y - plane and Γ is a line integral about this area, S_z is the z-component of the vector $\mathbf{S}$, $\mathbf{S}_t$ the orthogonal projection of $\mathbf{S}$ to the transverse plane and $\mathbf{n}$ is an outward normal unit vector from the contour Γ.

Using (2.8) and (2.9), and integrating over the infinite cross section leads to

$$\frac{\partial}{\partial z}P_\Omega(z) = \Re\left\{\int_\Omega i\epsilon_0\epsilon_r\omega \mathbf{E}\mathbf{E}^*\, d\Omega\right\} \tag{2.10}$$

$$P_\Omega(z) = \Re\left\{\int_\Omega S_z\, d\Omega\right\} \quad . \tag{2.11}$$

The line integral at infinity is identically zero because of the vanishing field at infinity. P_Ω indicates the whole power over a cross section, besides a factor of two. If the optical media are all loss-free, ϵ is purely real and the right hand side of (2.10) is identically zero. In this case it follows that we have $P_\Omega(z) = const$ for every cross section of the structure, as it is clear from the principle of the conservation of energy.

2.2. The Wave Equations for Structures with 1D Cross Sections

Because we are restricted to problems with 1D cross sections ϵ is set to be only a function of x and z. Additionally, the field solutions we look for should not change in the y-direction. With these assumptions we get the well known, independent TE (Transverse Electric) and TM (Transverse Magnetic) solutions of Maxwell's equations (2.1)–(2.4). From (2.5) and (2.2) it follows

$$\text{TE:} \quad \nabla^2 \mathbf{E}_y + n^2 k_0^2 \mathbf{E}_y = 0 \tag{2.12}$$

$$\mathbf{H}_x = -\frac{i}{\mu_0\omega}\frac{\partial}{\partial z}\mathbf{E}_y \tag{2.13}$$

$$\mathbf{H}_z = \frac{i}{\mu_0\omega}\frac{\partial}{\partial x}\mathbf{E}_y \quad , \tag{2.14}$$

and from (2.6) and (2.1)

$$\text{(2.15)} \qquad \text{TM:} \quad \nabla\left(\frac{1}{n^2}\nabla \mathrm{H}_y\right) + k_0^2 \mathrm{H}_y = 0$$

$$\text{(2.16)} \qquad \mathrm{E}_x = \frac{i}{\epsilon_0 \epsilon_r \omega}\frac{\partial}{\partial z}\mathrm{H}_y$$

$$\text{(2.17)} \qquad \mathrm{E}_z = -\frac{i}{\epsilon_0 \epsilon_r \omega}\frac{\partial}{\partial x}\mathrm{H}_y \quad ,$$

where we have introduced the refractive index n and the vacuum wave number k_0 by

$$\begin{aligned} n^2 &= \epsilon_r \\ k_0^2 &= \epsilon_0 \mu_0 \omega^2 \quad . \end{aligned}$$

The z-components of the appropriate Poynting vectors are given by

$$\begin{aligned} \text{TE:} \quad S_z \mathbf{e}_z &= -\mathrm{E}_y \times \mathrm{H}_x^* \\ &= -\frac{i}{\mu_0 \omega}\mathrm{E}_y \frac{\partial \mathrm{E}_y^*}{\partial z} \cdot \mathbf{e}_z \qquad \text{(2.18)} \\ \text{TM:} \quad S_z \mathbf{e}_z &= \mathrm{E}_x \times \mathrm{H}_y^* \\ &= \frac{i}{\epsilon \omega}\mathrm{H}_y^* \frac{\partial \mathrm{H}_y}{\partial z} \cdot \mathbf{e}_z \quad . \qquad \text{(2.19)} \end{aligned}$$

2.3. The Paraxial Wave Equations

As discussed above, we want to investigate a class of problems, which can be handled approximately as initial-boundary-value problems. For this purpose we use in the following the slowly varying amplitude or Fresnel approximation. The slowly varying amplitude $\tilde{E}_y$ is introduced by

$$\text{(2.20)} \qquad E_y = \tilde{E}_y e^{-in_0 k_0 x} \quad ,$$

where n_0 (real, positive) is not fixed so far.
Replacing E_y in (2.12) with (2.20) yields the transformed wave equation

$$\text{(2.21)} \qquad \frac{\partial^2 \tilde{E}_y}{\partial x^2} + \frac{\partial^2 \tilde{E}_y}{\partial z^2} + \left(n^2 - n_0^2\right) k_0^2 \tilde{E}_y = 2in_0 k_0 \frac{\partial \tilde{E}_y}{\partial z} \quad .$$

This equation is still exact, no approximation has been performed. If we are allowed to neglect the second term on the left hand side of (2.21), we get a

partial differential equation (PDE) with only a first derivation with respect to z. Clearly, this approximated PDE is a well defined initial-boundary-value problem, because the boundary conditions about the cross section are known to be the homogeneous Dirichlet conditions and a given field solution at $z = z_0$ uniquely determines every field distribution at $z \geq z_0$. Due to these reasons, the equation we want to solve for the TE-polarization is

$$\text{(2.22)} \qquad \text{TE:} \quad \frac{\partial^2 \tilde{E}_y}{\partial x^2} + \left(n^2 - n_0^2\right) k_0^2 \tilde{E}_y = 2in_0 k_0 \frac{\partial \tilde{E}_y}{\partial z} \quad .$$

The same argumentation holds for the equation (2.15) describing the TM-polarization, but in addition it is assumed for the moment that the refractive index is only a function of the transverse coordinate x. This leads to the approximated TM-equation

$$\text{(2.23)} \qquad \text{TM:} \quad \frac{\partial}{\partial x}\frac{1}{n^2}\frac{\partial \tilde{H}_y}{\partial x} + \frac{n^2 - n_0^2}{n^2} k_0^2 \tilde{H}_y = \frac{2in_0 k_0}{n^2}\frac{\partial \tilde{H}_y}{\partial z} \quad .$$

In practice equation (2.23) is used for slowly varying z-dependent refractive index changes too. This approximation is not examined here.
In the following the paraxial equations (2.22) and (2.23) are analyzed. To make this as easy as possible, the whole structure is assumed to be loss-free, i. e., the refractive index is purely real. The general idea is, that the power over a complete cross section (eq. (2.11)) remains constant throughout the structure. This follows from equation (2.10) and is a direct consequence from Maxwell's equations. But Maxwell's equations are violated by the paraxial wave equations (2.12) and (2.15), consequently the power calculated on the basis of the paraxial equations no longer remains constant. Therefore we can conclude two things:

- the deviation of the calculated power based on the paraxial wave equations from a constant value can be seen as a physical measure of the validity of these equations, and
- the free parameter n_0 should be chosen so that this deviation is as small as possible.

If the Poynting vector formulation equations (2.18) and (2.19) are rewritten in terms of the slowly varying amplitudes one gets

$$\text{(2.24)} \qquad \text{TE:} \quad S_z = -\frac{i}{\omega\mu_0}\tilde{E}_y\frac{\partial \tilde{E}_y^*}{\partial z} + \frac{n_0 k_0}{\omega\mu_0}\tilde{E}_y\tilde{E}_y^*$$

$$\text{(2.25)} \qquad \text{TM:} \quad S_z = \frac{i}{\omega\epsilon_0\epsilon_r}\tilde{H}_y^*\frac{\partial\tilde{H}_y}{\partial z} + \frac{n_0k_0}{\omega\epsilon_0\epsilon_r}\tilde{H}_y\tilde{H}_y^*$$

The case of the TE-polarization is considered first. To abbreviate the notation the usual L_2-scalar product of two complex functions $f_1(x)$ and $f_2(x)$ is introduced by

$$\text{(2.26)} \qquad (f_1, f_2) = \int_\Omega f_1(x)\cdot f_2^*(x)\,dx \quad .$$

The equations (2.11), (2.21), and (2.24) lead to a formula for the power P_Ω over a cross section, which is still exact:

$$\text{TE:} \quad P_\Omega = \frac{1}{2\omega\mu_0 n_0 k_0}\left\{-\left(\frac{\partial\tilde{E}_y}{\partial x},\frac{\partial\tilde{E}_y}{\partial x}\right) + k_0^2\cdot\left(\tilde{E}_y,(n^2-n_0^2)\tilde{E}_y\right) + \right.$$

$$\text{(2.27)} \qquad \left. + \left(\tilde{E}_y,\frac{\partial^2\tilde{E}_y}{\partial z^2}\right)\right\} + \frac{n_0k_0}{\omega\mu_0}\left(\tilde{E}_y,\tilde{E}_y\right) \quad .$$

If we step over to the paraxial approximation, the third power term on the right hand side of equation (2.27), containing the second z-derivative is neglected and P_Ω becomes inexact. To make one of the consequences clear, we assume for the moment that a z-independent waveguide is given. As it was shown in [8], the field distribution in every transverse plane of such a waveguide can be written as modal expansion of all transverse eigenfunctions

$$\text{(2.28)} \qquad \tilde{E}_y = \sum_j a_j(x)e^{-i\tilde{\beta}_j z} \quad ,$$

where forward and backward traveling waves are included. Therefore we have in the loss-free case (real β_j)

$$\text{(2.29)} \qquad \left(\tilde{E}_y,\frac{\partial^2\tilde{E}_y}{\partial z^2}\right) \le 0 \quad ,$$

independently from the direction of propagation. Thus, the calculated power based on the paraxial approximation is always larger than the true one, and we cannot decide if the power deviation is due to reflected waves or to the approximation of the forward propagating wave.
In order to keep P_Ω as accurate as possible, the neglected power term has to be much smaller than the other terms. As n_0 is the only parameter we have to influence the quality of our approximation, we require, that the power difference has to be a minimum in dependence of n_0.
Rewriting equation (2.27) leads to

$$\begin{aligned}\Delta P \cdot \omega\mu_0 &= \frac{1}{2n_0k_0}\left(\tilde{E}_y, \frac{\partial^2\tilde{E}_y}{\partial z^2}\right) \\ &= \omega\mu_0 P_\Omega - \frac{1}{2n_0k_0}T_1 - \frac{n_0k_0}{2}T_2 \quad , \end{aligned} \tag{2.30}$$

where T_1 and T_2 indicate the following n_0–independent terms

$$\begin{aligned} T_1 &= k_0^2\left(\tilde{E}_y, n^2\tilde{E}_y\right) - \left(\frac{\partial\tilde{E}_y}{\partial x}, \frac{\partial\tilde{E}_y}{\partial x}\right) \\ T_2 &= \left(\tilde{E}_y, \tilde{E}_y\right) \quad . \end{aligned}$$

For a local extremum we require

$$\begin{aligned}\frac{\partial\Delta P(n_0)}{\partial n_0} &= \frac{1}{2n_0^2k_0}T_1 - \frac{k_0}{2}T_2 \\ &= 0 \quad , \end{aligned}$$

ending up with

$$\begin{aligned} n_0^2 &= \frac{T_1}{k_0^2T_2} \\ &= \frac{k_0^2\left(\tilde{E}_y, n^2\tilde{E}_y\right) - \left(\frac{\partial\tilde{E}_y}{\partial x}, \frac{\partial\tilde{E}_y}{\partial x}\right)}{k_0^2\left(\tilde{E}_y, \tilde{E}_y\right)} \quad . \end{aligned} \tag{2.31}$$

This result is the same as proposed by SPLETT [13] to enlarge the applicability of Fresnel's wave equation. As we are looking only for positive n_0 and because T_2 is always positive too, T_1 has to be positive. Due to this fact, the second derivation of ΔP with respect to n_0 is always negative, i. e., a local maximum of the power difference ΔP was found. In Fig. 1 the qualitative dependence of the power difference on the refractive index n_0 is given.

If we use the original Helmholtz equation (2.12), we get a further interpretation of equation (2.31). Assuming a modal expansion for the original field of a z-uniform waveguide

$$E_y(x,z) = \sum_j a_j(x)e^{-i\beta_j z} \quad , \tag{2.32}$$

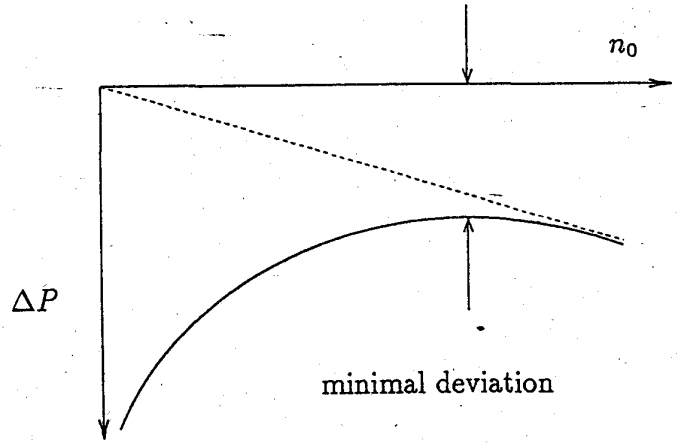


FIG. 1. Power deviation in dependence of the reference index

with a modal power P_j (see (2.18))

$$P_j = \frac{\beta_j}{\omega\mu_0} \cdot (a_j, a_j) \quad , \tag{2.33}$$

leads to

$$\begin{aligned} n_0^2 &= -\frac{\left(\frac{\partial^2 E_y}{\partial z^2}, E_y\right)}{k_0^2 (E_y, E_y)} \\ &= \sum_j \frac{\beta_j^2}{k_o^2} \frac{(a_j, a_j)}{(E_y, E_y)} \end{aligned} .$$

From this equation it is seen, that n_0 can be considered as the arithmetical mean of all propagation constants, weighted with their modal power, i. e., we take the first power weighted momentum of all propagation constants of the traveling wave group.
In summary, we can state the following properties of the paraxial wave equation (TE):

- there is an optimized approximation of the paraxial wave equation (2.22) to the original equation (2.21), if we compare them using the power difference equation (2.30)

- the reference index n_0 is given by equation (2.31). It is the same equation as holds for the propagation constant and the related field distri-

TABLE I. Normalization

	TE	TM
x	$k_0 x$	$k_0 x$
z	$k_0 z$	$k_0 z$
u	$\tilde{E}_y$	$\tilde{H}_y$
b	1	$\frac{1}{n^2}$
g	$n^2 - n_0^2$	$\frac{n^2 - n_0^2}{n^2}$
c	$2in_0$	$\frac{2in_0}{n^2}$
P	$\omega\mu_0 P_\Omega$	$\omega\epsilon_0 P_\Omega$

bution of a single eigensolution of a z-uniform waveguide.

- Suppose, the field distribution over a cross section of a z-uniform, loss-free structure is given. Then the power calculated using the paraxial approximation is always larger than the true power, except the case, where the field is given by a single eigensolution. In this special case the power difference vanishes.

The same argumentation as above given in detail for the TE-polarization holds for the paraxial TM-equation (2.23). In complete analogy we get

$$n_0^2 = \frac{k_0^2\left(\tilde{H}_y, \tilde{H}_y\right) - \left(\frac{1}{n^2}\frac{\partial \tilde{H}_y}{\partial x}, \frac{\partial \tilde{H}_y}{\partial x}\right)}{k_0^2\left(\frac{1}{n^2}\tilde{H}_y, \tilde{H}_y\right)} \tag{2.34}$$

2.4. NORMALIZATION

In order to simplify the notation, to avoid physical units and to handle both polarizations with the same equation, the normalization given in Table I is introduced.

In the following the notation of the left column of Tab. I is used.

Now the equation to solve is

$$\frac{\partial}{\partial x} b \frac{\partial}{\partial x} u + g \cdot u = c \cdot \frac{\partial}{\partial z} u \quad , \tag{2.35}$$

n_0 is given by

$$n_0^2 = \frac{(bn^2u, u) - \left(b\frac{\partial u}{\partial x}, \frac{\partial u}{\partial x}\right)}{(bu, u)} \quad , \tag{2.36}$$

and the power transmitted in the direction of propagation is

$$P = \frac{1}{2n_0}\left\{(gu, u) - \left(b\frac{\partial u}{\partial x}, \frac{\partial u}{\partial x}\right)\right\} + n_0(bu, u) \quad . \tag{2.37}$$

2.5. The Weak Formulation

In order to get an equivalent variational equation to (2.35) we use the standard procedure (see e. g. Zienkiewicz [16]). Equation (2.35) is multiplied with a function $v^*, v \in H_0^1$ and integrated over the cross section Ω

$$\int_\Omega \left(\frac{\partial}{\partial x} b \frac{\partial}{\partial x} u + g \cdot u - c \cdot \frac{\partial}{\partial z} u\right) \cdot v^*(x)\, dx = 0 \quad . \tag{2.38}$$

If the partial integration is performed, we get

$$a(u, v) = \left(c\frac{\partial u}{\partial z}, v\right) \quad , \tag{2.39}$$

where the bilinear form $a(u, v)$ is defined as

$$a(u, v) = (gu, v) - \left(b\frac{\partial u}{\partial x}, \frac{\partial v}{\partial x}\right) \quad . \tag{2.40}$$

Equation (2.39) is to solve for the weak solution $u \in H_0^1$ for all functions $v \in H_0^1$ and the homogeneous Dirichlet conditions $u = 0$ at Γ.
Using the bilinear form equation (2.40), the equations (2.36) and (2.37) can be rewritten as

$$\bar{n}_0^2 = \frac{a(u, u) + n_0^2(bu, u)}{(bu, u)} \quad , \tag{2.41}$$

and

$$P(u) = \frac{1}{2n_0} a(u, u) + n_0(bu, u) \quad , \tag{2.42}$$

where $\bar{n}_0$ indicates the refractive index according to equation (2.36). In general, n_0 and $\bar{n}_0$ will be differ, because the solution of the paraxial wave equation or the equivalent variational formulation depends on the original chosen refractive index n_0. But, comparing equation (2.36) to minimize the

power deviation and the definition of the bilinear form (2.40) it is found that the minimization condition is equivalently fulfilled, if

$$a(u,u) = 0 \quad . \tag{2.43}$$

In this case we have $\bar{n}_0 = n_0$ and additionally

$$P(u) = n_0(bu, u) \tag{2.44}$$

$$\left(c\frac{\partial u}{\partial z}, u\right) = 0 \quad . \tag{2.45}$$

Both equations are essential for the numerical algorithm. The first naturally leads one to use of the implicit midpoint rule as integration scheme, the second influences the performance of the longitudinal stepsize control.

3. The Numerical Algorithm

3.1. The Discretization in the Direction of Propagation

In order to get a fully discrete approximation of our basic equation (2.39), we follow Bornemann [2]-[3] and discretize *in the z-direction first.* It is assumed that we start the simulation at $z = 0$ with a given $u(x, o) = u_0$ and look for the solution $u(x, \Delta z) = u_{\Delta z}$, where Δz is the increment in the z-direction. Application of the midpoint rule to equation (2.39) leads to the stationary subproblem:

$$a_{\frac{\Delta z}{2}}\left(\frac{u_0 + u_{\Delta z}}{2}, v\right) = \left(c\frac{u_{\Delta z} - u_0}{\Delta z}, v\right)_{\frac{\Delta z}{2}} \quad , \tag{3.1}$$

where the subscript $\frac{\Delta z}{2}$ means that all z-dependent coefficients have to be taken at $z = \frac{\Delta z}{2}$. If the linear form $f(v)$ and the bilinear form $b(u, v)$ are introduced by

$$f(v) = \left(\left(c + \frac{\Delta z}{2} \cdot g\right) u_0, v\right)_{\frac{\Delta z}{2}} - \frac{\Delta z}{2}\left(b\frac{\partial u_0}{\partial x}, \frac{\partial v}{\partial x}\right)_{\frac{\Delta z}{2}} \quad , \tag{3.2}$$

and

$$b(u, v) = \left(\left(c - \frac{\Delta z}{2} \cdot g\right) u_0, v\right)_{\frac{\Delta z}{2}} + \frac{\Delta z}{2}\left(b\frac{\partial u_0}{\partial x}, \frac{\partial v}{\partial x}\right)_{\frac{\Delta z}{2}} \quad , \tag{3.3}$$

equation (3.1) can be rewritten as

$$b(u, v) = f(v) \quad . \tag{3.4}$$

3.2. The Complete Discretization

The subproblem (3.4) is solved in the usual finite element manner. A space $V_h \subset H_0^1$ is defined as

$$V_h = span\{v_1, \ldots, v_N\} \quad .$$

It is looked for an solution $u_h \in V_h$, which obeys

$$b(u_h, v_h) = f(v_h)$$

for all $v_h \in V_h$.
The whole cross section is divided into $N+1$ finite intervals as shown in Fig. 2.

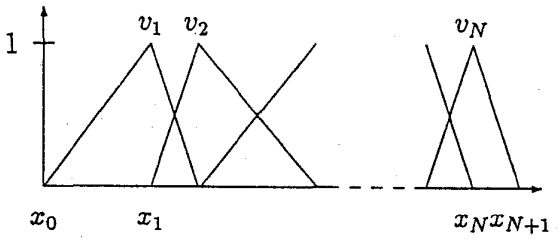


FIG. 2. Linear basis functions v_i

The functions v_i are given by

$$v_i = \begin{cases} v_{i-1}^R = \frac{1}{h_{i-1}}(x - x_{i-1}) & , \quad x_{i-1} \le x \le x_i \\ v_i^L = \frac{1}{h_i}(x_{i+1} - x) & , \quad x_i \le x \le x_{i+1} \\ 0 & , \quad elsewhere \end{cases} , \tag{3.5}$$

where

$$h_i = x_{i+1} - x_i \quad . \tag{3.6}$$

Every solution u_h, given by

$$u_h = \sum_{i=1}^{N} \alpha_i \cdot v_i$$

defines a vector u as

$$u = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix} \quad .$$

Now we can state the fully discrete problem as the solution of the linear system

$$\mathbf{B} \cdot \mathbf{u} = \mathbf{f} \quad , \tag{3.7}$$

where the matrix $\mathbf{B}$ and the vector $\mathbf{f}$ follow from their components

$$b_{ji} = b(v_i, v_j) \tag{3.8}$$

$$f_i = f(v_i) \quad . \tag{3.9}$$

3.3. The Problem of the Conservation of Energy

The problem of the conservation of energy will be discussed. We follow in principle the consideration outlined in STRANG/FIX [14] and show that using the implicit midpoint discretization and assuming the loss-free case, the optimized power evolution (2.45) is piecewise constantly approximated. If we write (3.1) again, using $v = (u_0 + u_{\Delta z})/2$, we get

$$\begin{aligned} a_{\frac{\Delta z}{2}}\left(\frac{u_0+u_{\Delta z}}{2}, \frac{u_0+u_{\Delta z}}{2}\right) &= \left(c\frac{u_{\Delta z}-u_0}{\Delta z}, \frac{u_0+u_{\Delta z}}{2}\right)_{\frac{\Delta z}{2}} \\ &= \left(\frac{c}{2\Delta z}u_{\Delta z}, u_{\Delta z}\right) - \left(\frac{c}{2\Delta z}u_0, u_0\right) + \\ &\quad + \left(\frac{c}{2\Delta z}u_{\Delta z}, u_0\right) - \left(\frac{c}{2\Delta z}u_0, u_{\Delta z}\right) \quad . \end{aligned}$$

Because the bilinear form $a(\cdot,\cdot)$ on the left hand side is purely real and c is purely imaginary (see Tab. I), we have

$$(c \cdot u_{\Delta z}, u_{\Delta z}) = (c \cdot u_0, u_0) \quad . \tag{3.10}$$

We see that in the discrete case the term

$$(c \cdot u, u) \tag{3.11}$$

is conserved along one step.

Additionally, viewing c as z-independent over one propagation step, which is exactly true for the TE-polarization and approximately valid for the TM-polarization, we can conclude from equation (2.45)

$$\begin{aligned} 0 &= \left(\frac{c}{i}\frac{\partial u}{\partial z}, u\right) \\ &= \Re\left\{\left(\frac{c}{i}\frac{\partial u}{\partial z}, u\right)\right\} \\ &= \frac{1}{2}\left[\left(\frac{c}{i}\frac{\partial u}{\partial z}, u\right) + \left(u, \frac{c}{i}\frac{\partial u}{\partial z}\right)\right] \\ &= \frac{1}{2}\frac{\partial}{\partial z}\left(\frac{c}{i}\cdot u, u\right) \quad , \end{aligned}$$

and therefore the continuous model requires that the term

$$\frac{1}{2}\left(\frac{c}{i}\cdot u, u\right) \tag{3.12}$$

is conserved within each propagation step. Using

$$c = 2in_0 b$$

introduced in Tab. I we find complete agreement with the power evolution equation (2.45). Comparing the equation (3.12) resulting from the continuous model and (3.11) from our discretization, we find that in fact (2.45) is approximated by a piecewise constant function. Another consequence from equation (3.11) is that (u, u) remains constant throughout the whole structure. As we approximate the optimal power evolution (2.45), based on Fresnel's model , any deviation from the constant power required by the complete Maxwell's equations allows to monitor the model error.

3.4. The Local Matrices

In order to calculate the local matrices we need the following results

$$\begin{aligned}
(v_i, v_i) &= \frac{1}{3}(h_{i-1} + h_i) \\
(v_i, v_{i+1}) &= \frac{1}{6}h_i \\
\left(\frac{\partial v_i}{\partial x}, \frac{\partial v_i}{\partial x}\right) &= \frac{1}{h_{i-1}} + \frac{1}{h_i} \\
\left(\frac{\partial v_i}{\partial x}, \frac{\partial v_{i+1}}{\partial x}\right) &= -\frac{1}{h_i} \quad .
\end{aligned}$$

Now we can decompose $\mathbf{B}$ into its local matrices $\mathbf{B}_i$ using the equations (3.3) and (3.8). As it will be explained later in the context of the adaptive refinement, the coefficients b, c, and g can set to be constant over a single finite element. Therefore we obtain

$$\mathbf{B}_i = \frac{h_i}{3}m_i\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} + \frac{s_i}{h_i}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{3.13}$$

for $i = 2 \ldots N-2$ and

$$\mathbf{B}_1 = \frac{h_1}{3}m_1\begin{pmatrix} 1+\frac{h_0}{h_1} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} + \frac{s_1}{h_1}\begin{pmatrix} 1+\frac{h_1}{h_0} & -1 \\ -1 & 1 \end{pmatrix} \tag{3.14}$$

$$(3.15)\mathbf{B}_{N-1} = \frac{h_{N-1}}{3} m_{N-1} \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 + \frac{h_N}{h_{N-1}} \end{pmatrix} + \frac{s_{N-1}}{h_{N-1}} \begin{pmatrix} 1 & -1 \\ -1 & 1 + \frac{h_{N-1}}{h_N} \end{pmatrix}$$

with

$$(3.16) \qquad m_i = c_i\left(\frac{\Delta z}{2}\right) - \frac{\Delta z}{2} g_i\left(\frac{\Delta z}{2}\right)$$

$$(3.17) \qquad \text{and} \quad s_i = \frac{\Delta z}{2} b_i\left(\frac{\Delta z}{2}\right) \quad .$$

Analogously, $\mathbf{f}$ from the right hand side of equation is decomposed into local vectors $\mathbf{f}_i$ using equation (3.2)

$$(3.18) \qquad f_i = \begin{pmatrix} m_i^0 \cdot \left(u_0, v_i^L\right) & - & s_i \cdot \left(\frac{\partial u_0}{\partial x}, \frac{\partial v_i^L}{\partial x}\right) \\ m_i^0 \cdot \left(u_0, v_i^R\right) & - & s_i \cdot \left(\frac{\partial u_0}{\partial x}, \frac{\partial v_i^R}{\partial x}\right) \end{pmatrix} \quad ,$$

where m_i^0 is defined as

$$(3.19) \qquad m_i^0 = c_i\left(\frac{\Delta z}{2}\right) + \frac{\Delta z}{2} g_i\left(\frac{\Delta z}{2}\right) \quad .$$

For $i = 1$, v_1^L has to be replaced by v_1 and for $i = N-1$, v_{N-1}^R has to be replaced by v_N.
Now the left hand side of the linear system (3.7) can be reformulated in terms of the local matrices (3.13) – (3.15)

$$(3.20) \qquad \mathbf{Bu} = \sum_{i=1}^{N-1} \mathbf{B}_i \left(u_i, u_{i+1}\right)^T$$

and for $\mathbf{f}$ one gets

$$(3.21) \quad \mathbf{f} = \left(f_1(1), f_1(2) + f_2(1), \ldots, f_i(2) + f_{i+1}(1), \ldots, f_{N-1}(2)\right)^T \quad .$$

The advantage of this formulation of the linear system (3.7) becomes obvious if we look at the adaptive refinement of a given grid, as we will see later.
The last question, which remains to answer, is how to approximate numerically the scalar products contained in equation (3.18). Concerning the practical problems discussed in this paper, the use of the two point Gauss quadrature turned out to be a good choice. With the appropriate formula we get for the first type of the scalar products:

$$\begin{aligned}\int_0^h u_0(x)\cdot v^L(x)\,dx &\approx \frac{h}{2}\left(\omega_1 u_0(x_1)+\omega_2 u_0(x_2)\right)\\ \int_0^h u_0(x)\cdot v^R(x)\,dx &\approx \frac{h}{2}\left(\omega_2 u_0(x_1)+\omega_1 u_0(x_2)\right)\end{aligned}\quad,$$

where the constants are given by

$$\begin{aligned}\omega_1 &= \frac{1}{2}+\frac{\sqrt{3}}{6}\\ \omega_2 &= \frac{1}{2}-\frac{\sqrt{3}}{6}\\ x_1 &= \frac{h}{2}\left(1-\frac{\sqrt{3}}{6}\right)\\ x_2 &= \frac{h}{2}\left(1+\frac{\sqrt{3}}{6}\right)\end{aligned}\quad.$$

The remaining type of scalar products can be integrated exactly:

$$\begin{aligned}\int_0^h \frac{\partial u_0}{\partial x}\cdot\frac{\partial v^L}{\partial x}\,dx &= -\frac{1}{h}\left(u_0(h)-u_0(0)\right)\\ \int_0^h \frac{\partial u_0}{\partial x}\cdot\frac{\partial v^R}{\partial x}\,dx &= \frac{1}{h}\left(u_0(h)-u_0(0)\right)\end{aligned}\quad.$$

3.5. Adaptive Refinement of the Finite Element Grid

For every given cross section, the stationary subproblem (3.4) has to be solved up to a prescribed tolerance η_t. The first problem is the definition of a fundamental grid, from which the refinement is started. The class of integrated optics devices, which is investigated in this paper, is characterized by the fact, that the whole component is composed by a few number of thin films. The refractive index is constant within each layer, but changes between two neighboured films (see Fig. 3). Due to this refractive index geometry we can define a fundamental mesh in such a way that every single layer is described by a single finite element. The so chosen grid at level zero describes the complete transverse index structure but contains a low number of nodes.

On the fundamental grid a solution is computed, which is of course very rough. In order to select the elements to subdivide, we need an estimator

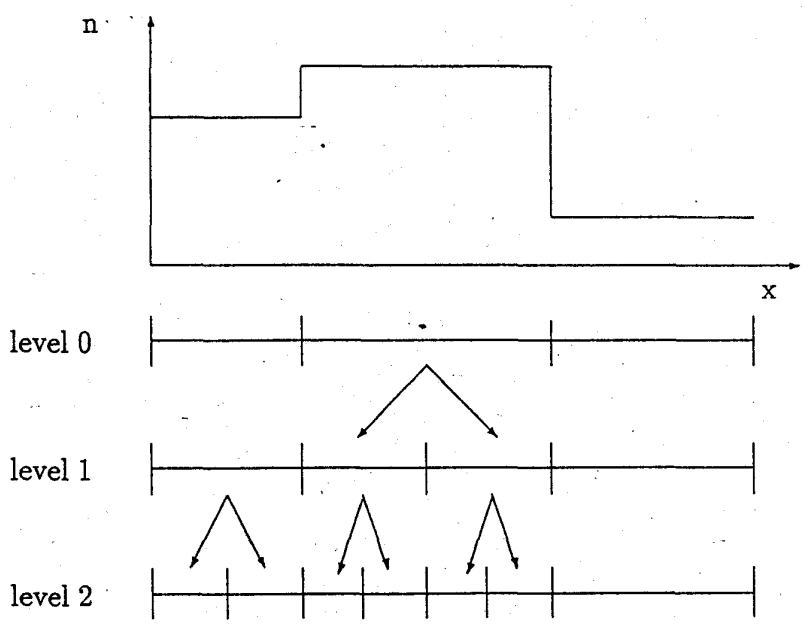


FIG. 3. Mesh development

for the local errors belonging to every finite element and an appropriate refinement strategy. The local error estimator used was proposed in [2] and results from a comparison between the solution obtained using linear finite elements and a local improvement using quadratic elements. Equation (3.4), written only for a single finite element reads

$$m_i \cdot (u,v) + s_i \cdot \left(\frac{\partial u}{\partial x}, \frac{\partial v}{\partial x}\right) = m_i^0 \cdot (u_0, v) - s_i \cdot \left(\frac{\partial u_0}{\partial x}, \frac{\partial v}{\partial x}\right) \quad . \tag{3.22}$$

It is assumed that the linear solution u^l is known and we look for a quadratic improvement u^q. We introduce

$$u = u^l + u^q \tag{3.23}$$

$$u^q = \alpha \cdot v^q \tag{3.24}$$

$$v^q = -\frac{4}{h^2} x\,(x - h_i) \quad . \tag{3.25}$$

Inserting equation (3.23) – (3.25) into equation (3.22) leads to

$$\alpha = \frac{m^0(u_0, v^q) - m(u^l, v^q) - s\left(\frac{\partial u_0}{\partial x}, \frac{\partial v^q}{\partial x}\right) - s\left(\frac{\partial u^l}{\partial x}, \frac{\partial v^q}{\partial x}\right)}{m\,(v^q, v^q) + s\,(v_x^q, vq_x)} \quad . \tag{3.26}$$

The evaluation of the scalar products gives

$$\begin{aligned}
\left(u^l, v^q\right) &= \frac{h}{3}\left(u^l(0) + u^l(h)\right) \\
\left(v^q, v^q\right) &= \frac{8}{15}h \\
\left(\frac{\partial v^q}{\partial x}, \frac{\partial v^q}{\partial x}\right) &= \frac{16}{3h} \\
\left(\frac{\partial u^l}{\partial x}, \frac{\partial v^q}{\partial x}\right) &= 0 \\
\left(u_0, v^q\right) &\approx \frac{h}{9}\left(u(x_1) + 4u(x_2) + u(x_3)\right)
\end{aligned}$$

with

$$\begin{aligned}
x_1 &= h \cdot \left(\frac{1}{2} - \frac{\sqrt{15}}{10}\right) \\
x_2 &= \frac{h}{2} \\
x_3 &= h \cdot \left(\frac{1}{2} + \frac{\sqrt{15}}{10}\right) \quad ,
\end{aligned}$$

and

$$\begin{aligned}
\left(\frac{\partial u_0}{\partial x}, \frac{\partial v^q}{\partial x}\right) &\approx \frac{4}{h}\left(-u(0) + u(x_1) + u(x_2) - u(h)\right) \\
x_1 &= h \cdot \left(\frac{1}{2} - \frac{\sqrt{3}}{6}\right) \\
x_2 &= h \cdot \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right) \quad .
\end{aligned}$$

The approximation of the last two scalar products was performed using the three point Gauss quadrature. With α from equation (3.26) the estimated transversal error of the ith finite element is

$$\begin{aligned}
\eta_{t_i} &= \sqrt{(\alpha v^q, \alpha v^q)} \\
&= |\alpha| \sqrt{\frac{8}{15}h} \quad .
\end{aligned}$$

Consequently, the resulting error estimates is given by

$$\eta_t = \sqrt{\sum_{i=1}^{N} \eta_{t_i}^2} \quad .$$

Now we are able to build a refinement strategy:

refine the i-th element if $\eta_i > cut$.

To determine 'cut', we use a simple prediction scheme following BABŮSKA and RHEINBOLDT [1] to forecast what may be happen to η_i if the belonging element is subdivided. Locally we may assume that the following representation of the local error indicator is valid

$$\eta_i = c_i h_i^{\lambda_i} \quad .$$

Taking the history of the refinement into account, we also have approximately

$$\eta_i^{old} = c_i (2h_i)^{\lambda_i}$$

and

$$\eta_i^{new} = c_i \left(\frac{h_i}{2}\right)^{\lambda_i} \quad .$$

and therefore the prediction η^{new} for the new elements, if the actual elements will be subdivided, is estimated to

$$\eta_i^{new} = \frac{\eta_i^2}{\eta^{old}} \quad .$$

Obviously, only those elements should be subdivided, which have an η_i-value above the largest predicted new η-value.

$$cut := \max_i \eta_i^{new}$$

3.6. THE LONGITUDINAL STEPSIZE CONTROL

Until this stage, we have achieved the following:

- The z-discretization was carried out first, using a physically reasonable stepsize Δz

- The arising stationary subproblem was solved adaptively using the finite element method. Any prescribed accuracy concerning this subproblem can be achieved.

Since we want to limit the complete discretization error rather than only the transversal one, the following requirement has to be fulfilled:

$$\eta_t + \eta_z \leq TOL \quad , \tag{3.27}$$

where η_z indicates the longitudinal error still required and TOL is the given accuracy. In order to obtain η_z , we compare the solution u of equation (3.1) and a new solution u^e , which results from a implicit Euler discretization of the basic equation (2.39). The implicit Euler step can be written as

$$a_{\Delta z}\left(u^e_{\Delta z}, v\right) = \left(c(\Delta z) \cdot \frac{u^e_{\Delta z} - u^e_0}{\Delta z}, v\right) \quad .$$

In complete analogy to the sections 3.2 and 3.4 we get a new linear system

$$\mathbf{B}^e \mathbf{u}^e = \mathbf{f}^e \quad . \tag{3.28}$$

The matrix $\mathbf{B}^e$ can be assembled from the local matrices (3.13) – (3.15) exactly as $\mathbf{B}$ does, but the factors m_i, m_i^0 , and s_i have to be changed to m_i^e, $m_i^{0,e}$, and s_i^e by

$$\begin{aligned} m_i^e &= c_i(\Delta z) - \Delta z \cdot g_i(\Delta z) \\ m_i^{0,e} &= c_i(\Delta z) \\ s_i^e &= \Delta z \cdot b_i(\Delta z) \quad . \end{aligned}$$

The same holds for the local vectors $\mathbf{f}_i^e$, which form the right hand side of (3.28). We obtain $\mathbf{f}_i^e$ from $\mathbf{f}_i$ of (3.18), if there s_i is set to be zero and m_i^0 is replaced again by $m_i^{0,e}$.
Now η_z is given by

$$\eta_z = \|u - u^e\|_{L_2} \quad .$$

Following DEUFLHARD [4], [5], a proposal for a new stepsize can be obtained by

$$\Delta z^{new} = \sigma_z \sqrt{\frac{\frac{1}{2}TOL}{\eta_z}} \Delta z \quad . \tag{3.29}$$

where σ_z is a constant. Now the complete adaptive algorithm can be described:

1. For the proposed Δz the stationary subproblem is solved with an accuracy η_t, where
$$\eta_t \leq \sigma_t \cdot \frac{1}{2} \cdot TOL$$
is required. In our examples a security factor $\sigma_t = 0.25$ was chosen.

2. Next, the same step on the same grid is computed again, using now the implicit Euler discretization and η_z is calculated.

3. According to equation (3.29), Δz^{new} is determined. If is the resulting discretization error is larger than TOL, then the whole propagation step has to be calculated again using the proposed new Δz, otherwise the next step can be calculated starting with $\Delta z = \Delta z^{new}$.

TABLE II. Input waveguide

waveguide geometry		refractive index n	
width	$0.4\mu m$	substrate	3.16446
length	$1000\mu m$	waveguide	3.39885

4. APPLICATION TO TAPER STRUCTURES

The cause to develop the adaptive algorithm described in the previous sections was the need for an effective numerical tool to simulate integrated optical waveguide tapers with complicated refractive index geometries based on the semiconductor material InGaAsP/InP. Such taper structures can be seen as mode transforming devices connecting dielectric waveguides with different spot sizes. As an example, the coupling of a passive semiconductor waveguide to a glass fiber will be investigated in this section. The structure to simulate has been proposed by Nolting [9] as a part of the COST-240 problem set. The spatial refractive index distribution is compatible with present fabrication technologies. In order to learn some properties of the adaptive algorithm, we will build up the 1D-model of the original structure in several steps. At first only the input and output waveguides of the structure are considered, then the case of the butt coupling of both waveguides, followed by the case of a symmetrical tapering, and finally the complete taper model. The program AMIO1 (mnemotechnically for: Adaptive Multilevel Integrated Optics 1–D) is written in the language C. All simulations presented in this section were carried out on a SUN sparc workstation IPC.

4.1. THE Z-UNIFORM SLAB WAVEGUIDES

Figure 4 shows the principal spatial refractive index geometry of the two incorporated z-uniform waveguides and their data are given in Table II and III respectively.

If the input waveguide is launched with its fundamental mode at $z = 0$, the fields propagates in the z-direction and the local amplitude distribution $|u|$ remains unchanged as it is shown in Figure 5. The distribution of nodes belonging to this field distribution is given in Figure 6. The concentration of nodes is higher in regions with higher field intensities and becomes low far away from the waveguide. The simulation of the output waveguide leads to very similar plots and is therefore suppressed here. The performance of both

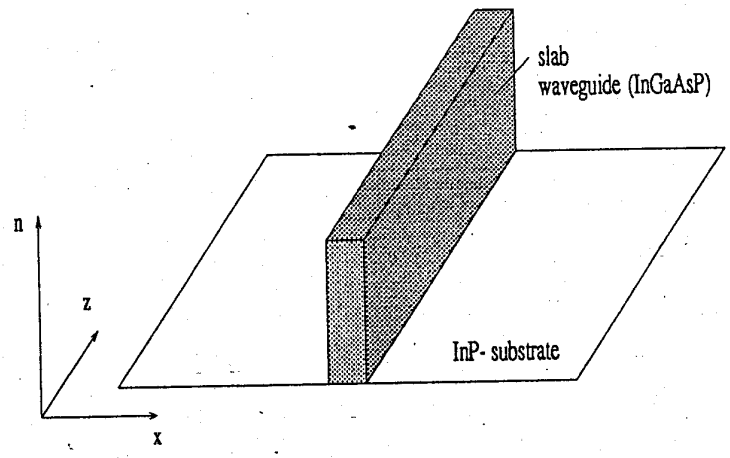


FIG. 4. Z-uniform slab waveguide

TABLE III. Output waveguide

waveguide geometry		refractive index n	
width	$5.0\mu m$	substrate	3.16446
length	$1000\mu m$	waveguide	3.16756

simulation runs is given in Table IV.

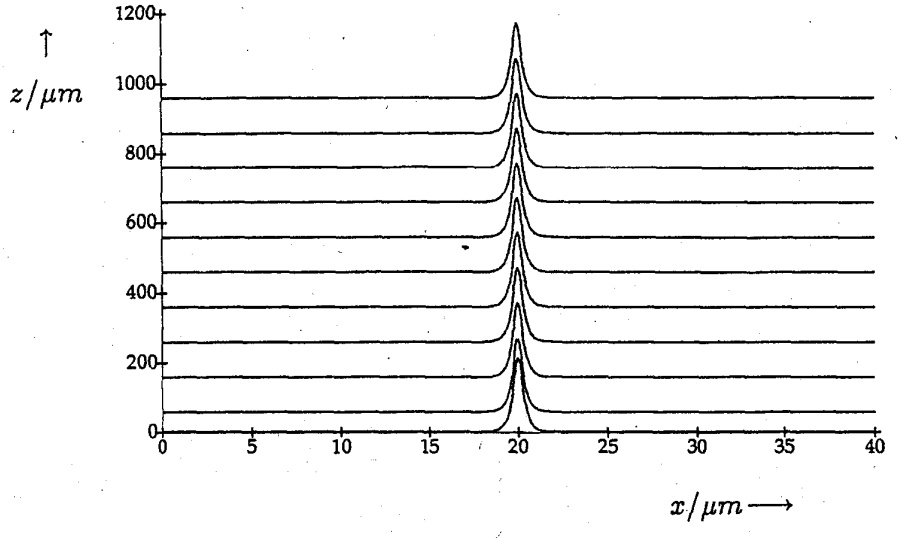


FIG. 5. Z-uniform slab waveguide: field distribution

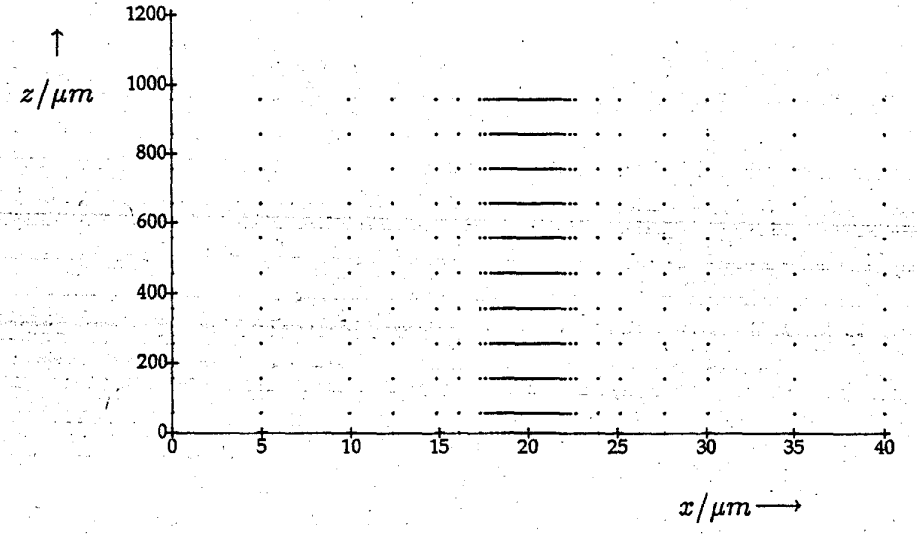


FIG. 6. Z-uniform slab waveguide: self-adaptive distribution of nodes of Figure 5

Because the analytical solution u_{true} of the field propagating along the z-uniform waveguide is known, we can compute the true error

TABLE IV. Numerical effort

	TOL	CPU/s	steps
input waveguide	0.01	17.3	54
output waveguide	0.01	14.0	53

$$\eta_{true} = \sqrt{(u - u_{true}, u - u_{true})} \quad .$$

The appropriate results are plotted in the Figures 7 and 8.

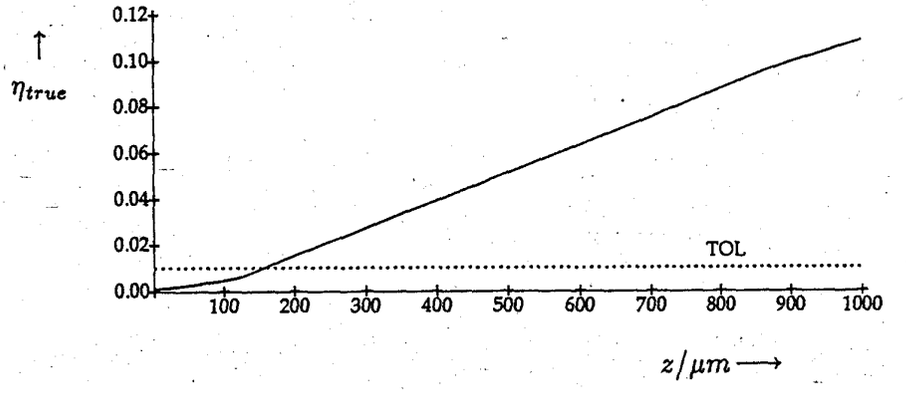


FIG. 7. Input waveguide: true error

In both cases the error increases linearly after a short transient length, but the magnitude of the errors is very different. Whereas the error of the weak guiding output waveguide remains less than the prescribed TOL along the whole simulation length, the error of the strong guiding waveguide exceeds this value after $150\mu m$. These errors are due to a deviation between the exact propagation constant β and the numerical generated propagation constant β_{num}. Of course the difference $\beta - \beta_{num}$ can be decreased by decreasing the prescribed tolerance TOL, but this leads to an increasing numerical effort and seems to be not necessary for most of the practical purposes. If we compute the the accumulated phase differences along the whole simulation length $\Delta\phi = (\beta - \beta_{num}) \cdot z$, we get for the strong guiding input waveguide

$$\Delta\phi = 6.8^0 \quad (0.12rad)$$

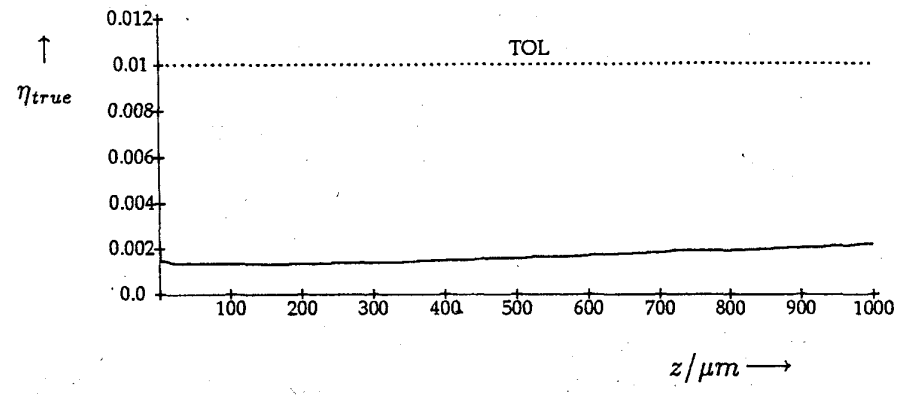


FIG. 8. Output waveguide: true error

and for the weak guiding output waveguide

$$\Delta\phi = 0.1^0 \quad (0.1710^{-3} rad) \quad .$$

Concerning practical cases, where relative phase differences play a more essential role than absolute phase differences, one is encouraged to start a simulation with even larger tolerances.

4.2. THE BUTT COUPLING

Figure 9 shows the butt coupling of the input and the output waveguide considered in the previous section. The length of both waveguides is taken to $50\mu m$. The field distribution of this structure, again launched by the fundamental mode of the input waveguide, is given in Figure 10 and the belonging distribution of nodes in Figure 11. These figures show exemplarily the adaptive discretization of the transverse cross section and the influence of the stepsize control. A very small number of nodes and steps is needed to propagate the field along the input waveguide, whereas this number dramatically increases behind the location of the butt coupling (see Figure 12). The numerical characteristics are given in Table V.

As discussed in section above, the power $P(z)$ can be used to estimate the validity of Fresnel's approximation to Maxwell's equations. Figure 13 shows the plots of the power $P(z)$ and the scalar product (u, u) along the direction of transmission. The scalar product remains constant due to the choice of the implicit midpoint rule, but the power changes abruptly at the location of

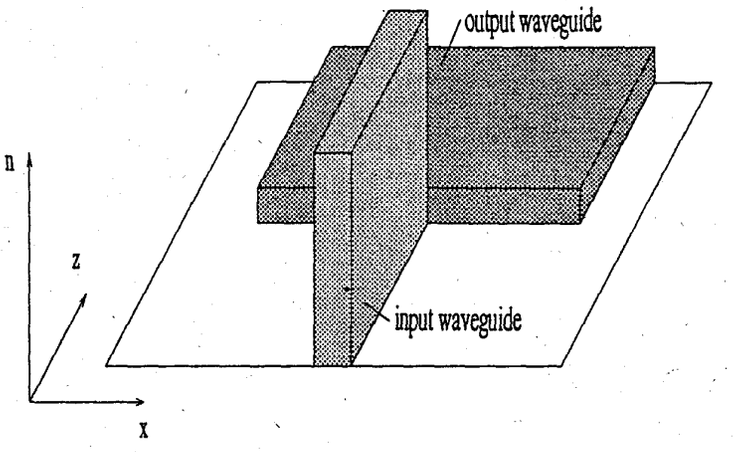


FIG. 9. Butt coupling of the input and the output waveguide

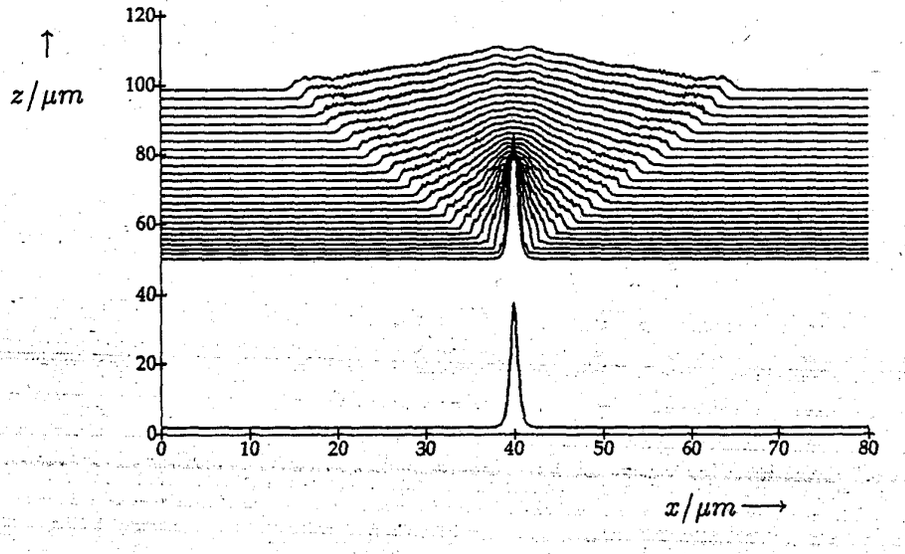


FIG. 10. Butt coupling: field distribution

TABLE V. Numerical effort to simulate the butt coupling

TOL	CPU/s	steps 1. section	steps 2. section
0.05	157.0	7	147

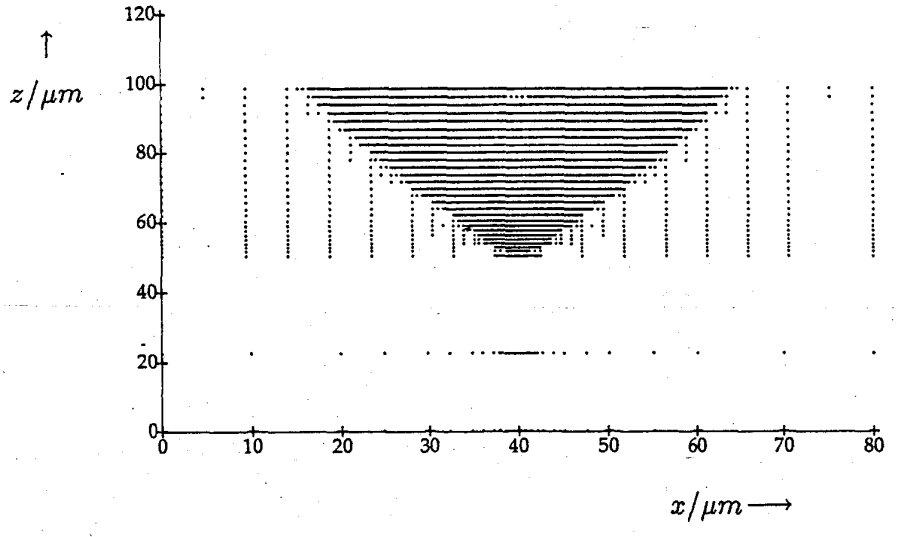


FIG. 11. Butt coupling: distribution of nodes

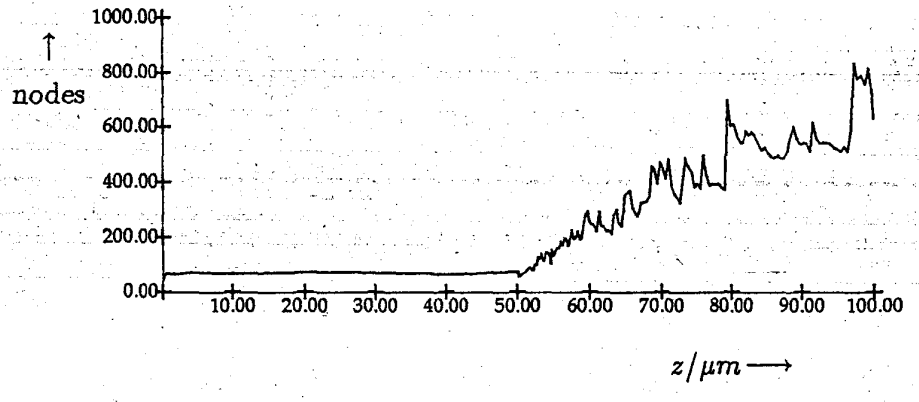


FIG. 12. Butt coupling: evolution of the number of nodes

TABLE VI. Geometry of the symmetrical taper structure

width at the input face	$0.4\mu m$
width at the output face	$0.01\mu m$
length of the input guide	$50.0\mu m$
length of the output guide	$50.0\mu m$
taperlength	$500.0\mu m$

the refractive index change. Because an ideal solution of a nonlossy structure is characterized by $P = const.$, every power deviation from a constant value indicates an inaccuracy of our model based on Fresnel's wave equation. From this point of view we can explain the abrupt power change with the generation of reflected waves, which are not taken into account in our propagation model.

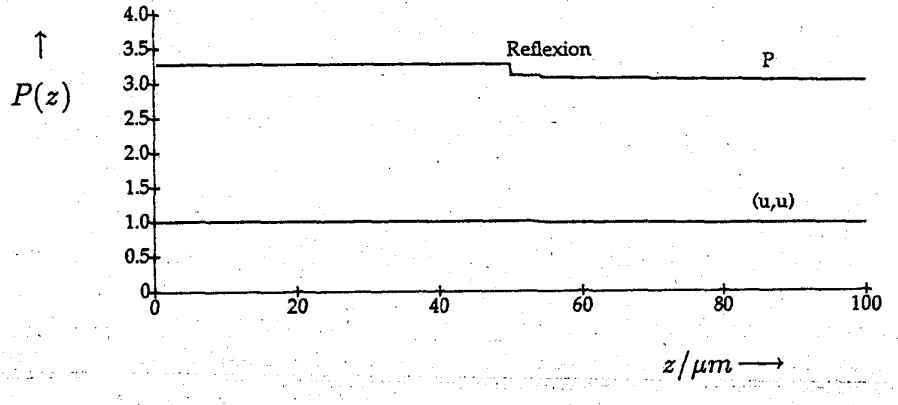


FIG. 13. Butt coupling: power evolution, behavior of P allows monitoring of the modeling error

4.3. THE SYMMETRICAL TAPER STRUCTURE

The principal spatial refractive index structure of the symmetrical taper structure is shown in Figure 14, the related data are given in Table VI.

The taper itself has the same refractive index as the input waveguide and its width is linearly decreased. This structure is well suited to analyze some properties of the numerical algorithm. In order to compare the different results using different tolerances, the definition of the overlap integral is

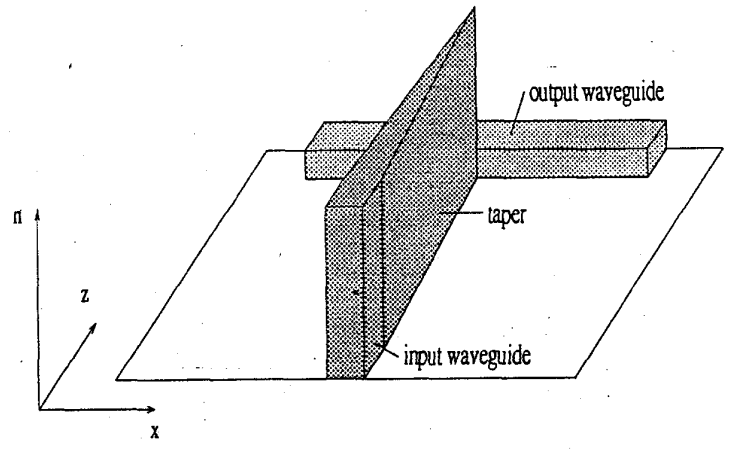


FIG. 14. Symmetrical taper

useful. We assume that we can expand every solution $u(z)$ in a series of modal eigenfunctions $a_i(x)$ of a z-uniform reference waveguide.

$$u(x) = \sum_i \alpha_i a_i(x) \quad ,$$

with the normalization $(a_i, a_i) = 1$. Using the orthogonality relation

$$(a_i, a_j) = 0 \qquad \text{for} \qquad i \neq j \quad ,$$

which is valid in the nonlossy case, we get

$$\alpha_i = (u(z), a_i) \quad .$$

Now, $|\alpha_0^2|$ provides an approximated measure for the power coupling efficiency from the input to the fundamental mode of the reference waveguide. This can be seen as follows. If the input waveguide is launched with its fundamental mode a_0^{in} obeying the normalization condition, and having a propagation constant $\beta_0^{in} = n_0^{in} \cdot k_0$, then this mode carries a power, according to equation (2.45)

$$P^{in} = n_0^{in} \quad .$$

Equivalently, we get the power of the fundamental mode of any reference waveguide

TABLE VII. Numerical effort and results using different tolerances

TOL	CPU/s	steps	η
0.08	19.3	72	0.8654
0.06	23.2	84	0.8651
0.04	30.2	105	0.8686
0.02	60.9	160	0.8621
0.01	124.6	254	0.8640

$$\begin{aligned} P^{ref} &= n_0^{ref} \cdot \left(\alpha_0^{ref} a_0^{ref}, \alpha_0^{ref} a_0^{ref}\right) \\ &= n_0^{ref} \cdot \left|\alpha_0^{ref}\right|^2 \quad . \end{aligned}$$

Now, the coupling efficiency is

$$\begin{aligned} \eta &= \frac{P^{ref}}{P^{in}} \\ &= \frac{n_0^{ref}}{n_0^{in}} \cdot \left|\alpha_0^{ref}\right|^2 \quad ; \end{aligned}$$

For practical purposes the last equation is approximated by

$$\eta = \left|\alpha_0^{ref}\right|^2 \quad .$$

This equation is often referred as overlap integral.
The symmetrical taper structure was simulated varying the tolerance from 0.01 to 0.08. Table VII shows the numerical effort needed and the obtained coupling coefficients η, if the output waveguide is taken as reference waveguide.

The η-values differ less than 1 per cent from each other, i. e., even the rough tolerance $TOL = 0.08$ is sufficient to calculate the power coupling effieciency.
To give an idea, how the computed field behaves using different tolerances, the field distributions and their distributions of nodes are given for $TOL = 0.08$ (Figures 15 and 16) and for $TOL = 0.01$ (Figures 17 and 18). The more accurate simulation offers some more details than the coarser one. To compare both simulations more quantitatively, the overlap integral according to equation (4.3.) is used in a modified way. Along the input waveguide

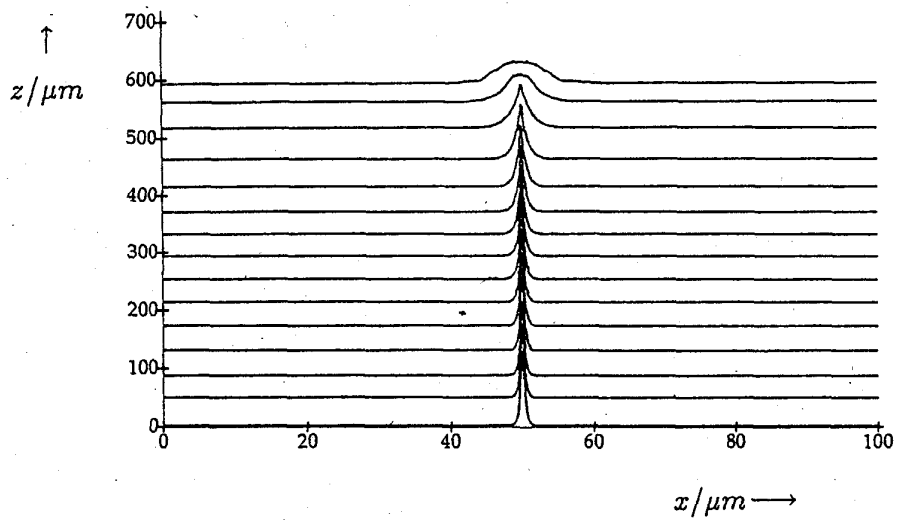


FIG. 15. Symmetrical taper: field distribution, $TOL = 0.08$ (compare Figure 20)

$\eta(z)$ is calculated using the input waveguide itself as reference waveguide to compute $a_0(x)$, i. e., the normalized input power is observed. Behind this waveguide the output waveguide is taken as reference waveguide, i. e., now $\eta(z)$ gives the power coupling coefficient to the output waveguide. Figure 19 shows $\eta(z)$ computed in the described manner. One can read from this figure that without the taper the power coupling coefficient will be about 0.26 and that this value is increased using the taper up to 0.86. The plot of the different results $\eta(z)$, one using $TOL = 0.01$ and the other using $TOL = 0.08$ in the same coordinate system cannot be distinguished from each other.

All results presented so far used a reference index n_0 according to equation (2.36). The intuitive feeling that this choice of n_0 not only guarantees a minimum of the power error but also seems to be preferable with respect to the stepsize control is caused by the physical insight that we have transformed the reference index to the first power weighted momentum of all propagation constants of the launched wave group and that therefore the algorithm sees a minimum of changes in the direction of propagation. This is underlined by equation (2.45). To prove this, Figure 20 shows the simulation of the symmetrical taper again, using $TOL = 0.08$ and taking the reference index to the refractive index of the InP-substrate as it is done usually. Comparing the Figures 15 and 20 the shorter steplength in the lower part of the figure is

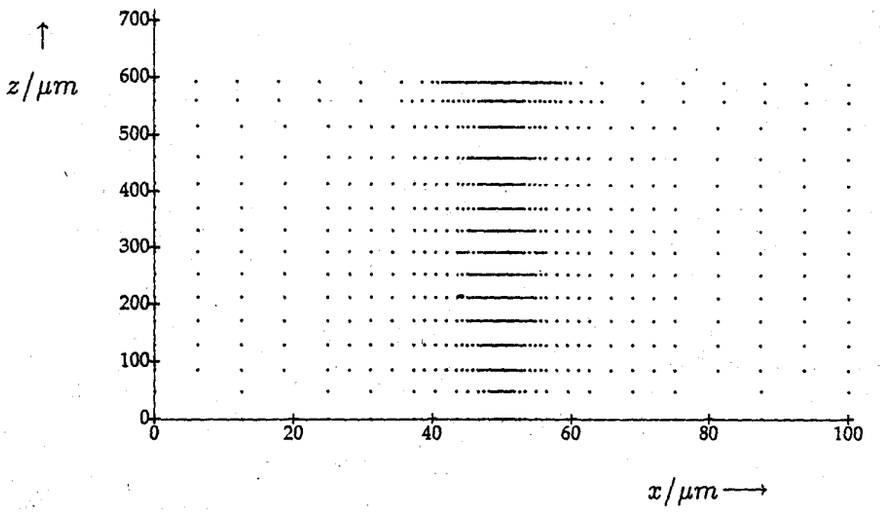


FIG. 16. Symmetrical taper: distribution of nodes, $TOL = 0.08$

TABLE VIII. Numerical effort using different reference indexes

	TOL	CPU/s	steps
adap. n_0	0.08	19.3	72
$n_0(InP)$	0.08	104.8	468

visible. The increasing stepsize in Figure 20 can be explained with fact that the difference between the substrate index and the reference index according equation (2.36) gradually decreases along the taper structure (Figure 21). Table VIII gives a comparison between both simulations.

The Figures 22 and 23 prove that the number of nodes is not essentially influenced by the choice of n_0.

The last representation using the symmetrical taper structure concernes the integration scheme for the z-integration. As discussed in section 3.3 the choice of the implicit midpoint rule seems to be well suited to approximate the required conservation of power. In fact, Figure 24 comparing the power evolution using the midpoint rule in a first run and the Crank-Nicolson scheme in a second run shows that the required power constancy is significantly better approximated applying the midpoint rule.

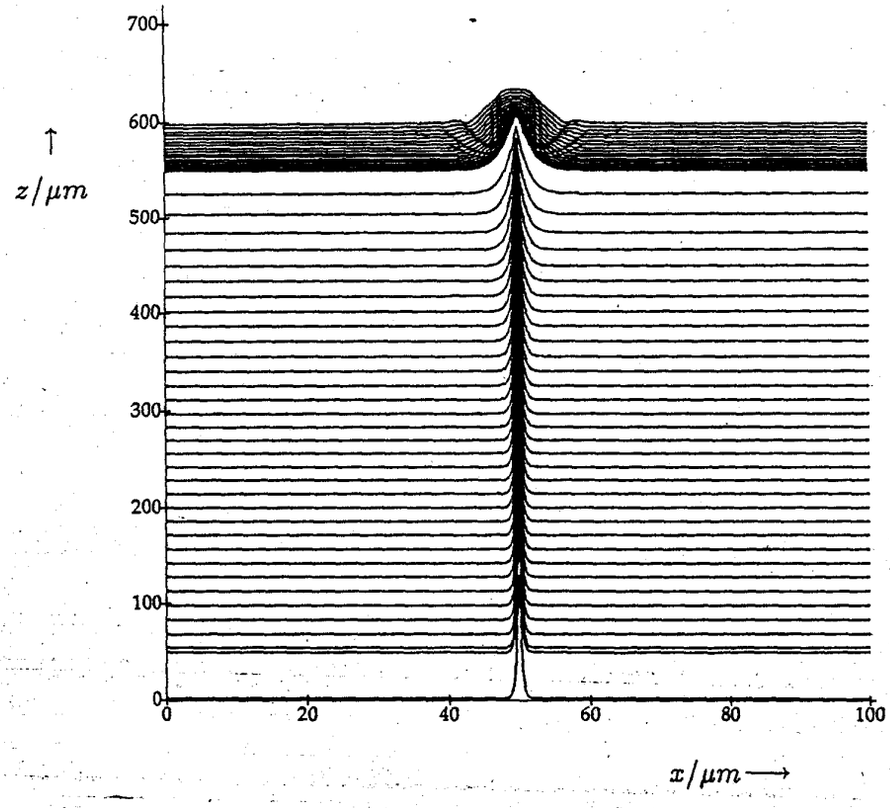


FIG. 17. Symmetrical taper: fielddistribution, $TOL = 0.01$

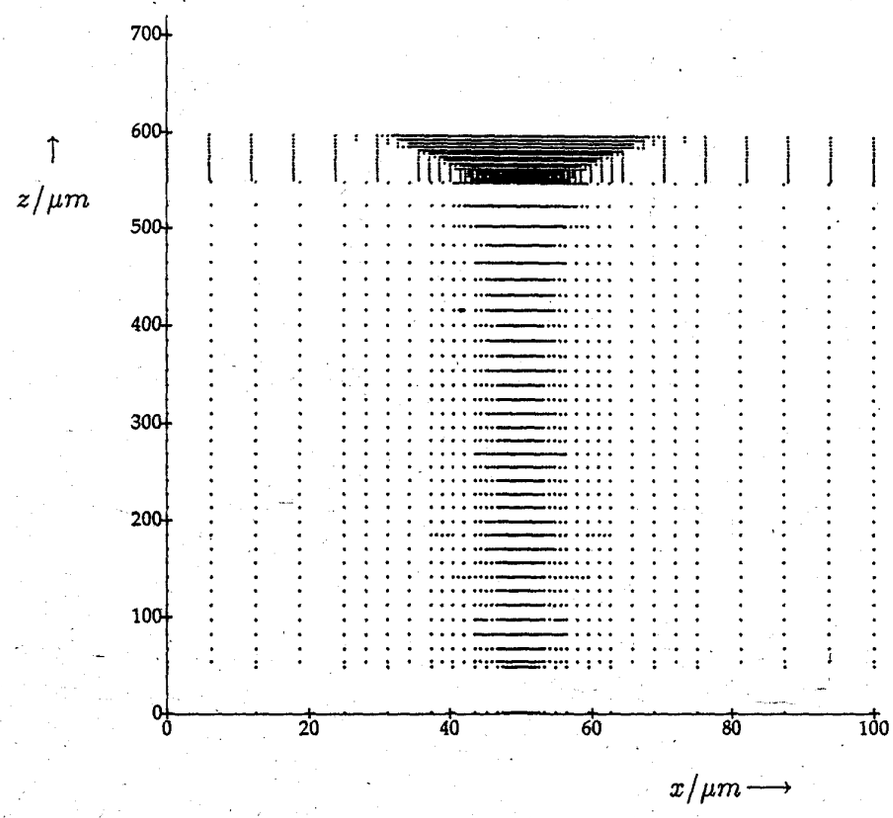


FIG. 18. Symmetrical taper: distribution of nodes, $TOL = 0.01$

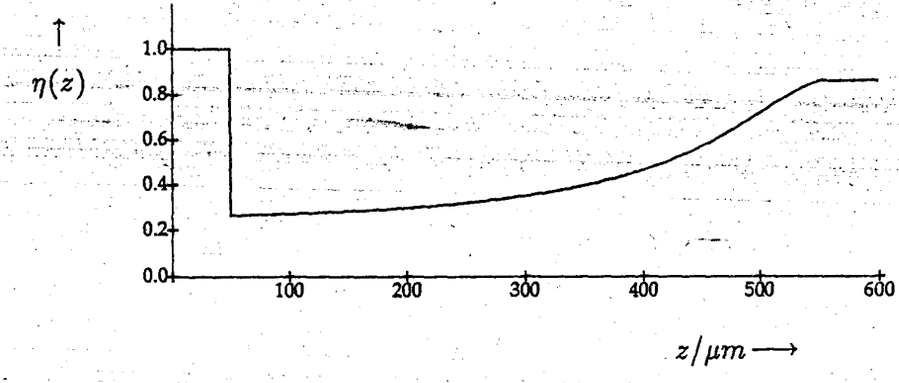


FIG. 19. Symmetrical taper: power coupling coefficient

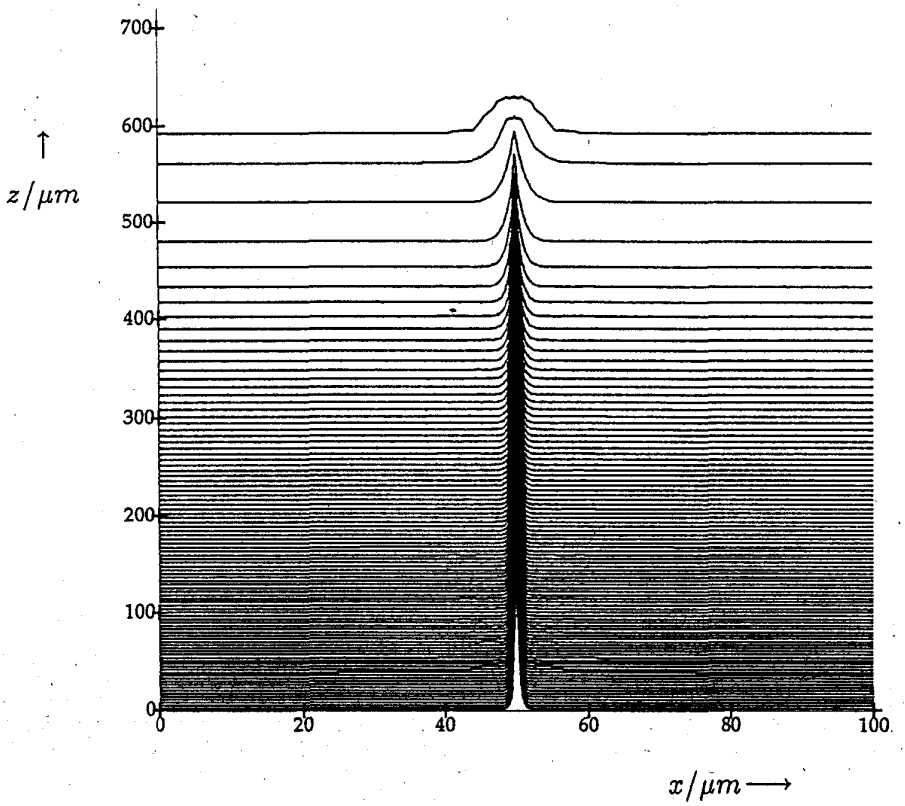


FIG. 20. Symmetrical taper: field distribution, fixed $n_0 = n(InP)$, $TOL = 0.08$ (compare Figure 15)

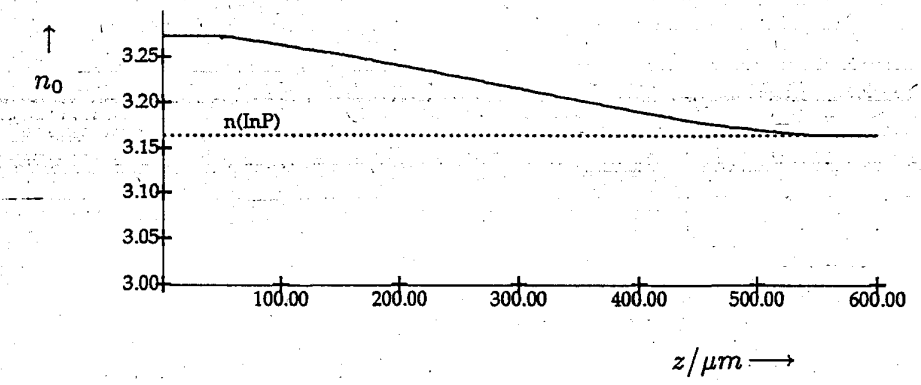


FIG. 21. Symmetrical taper: adaption of n_0

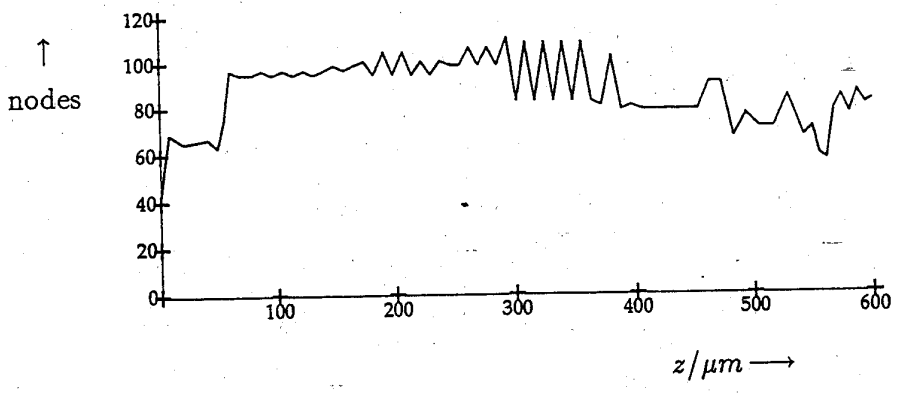


FIG. 22. Number of nodes using the adaptive n_0

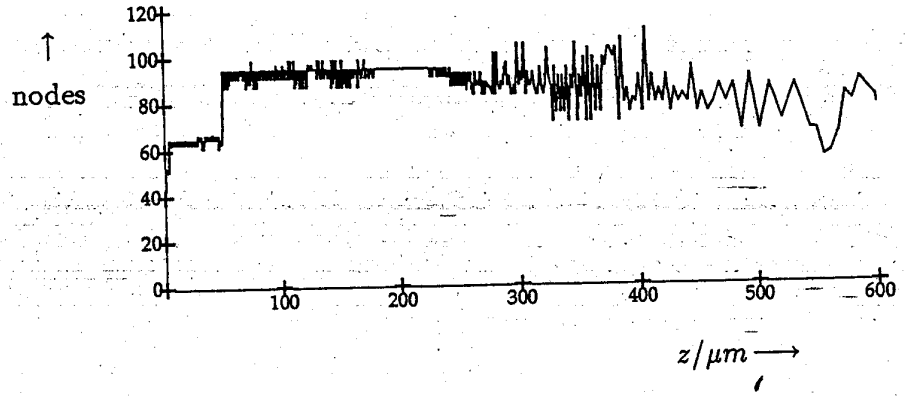


FIG. 23. Number of nodes using the fixed $n_0(InP)$

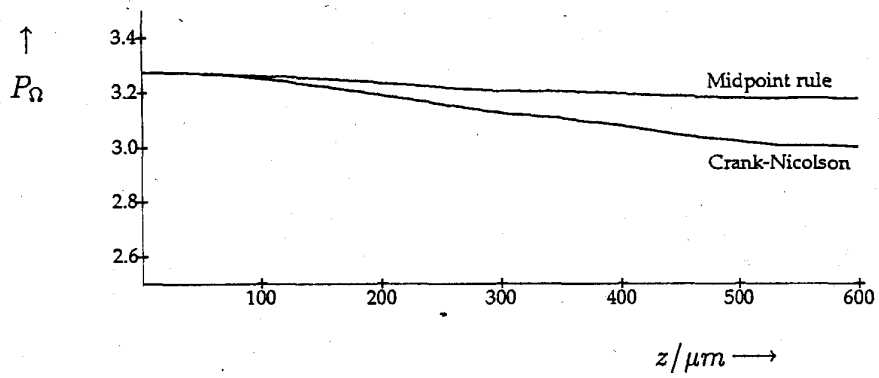


FIG. 24. Symmetrical taper: Comparison of the power deviation using Crank-Nicolson and implicit midpoint integration scheme

TABLE IX. Geometry of the asymmetrical taper structure

width at the input face	$0.4\mu m$
width at the output face	$0.01\mu m$
length of the input guide	$50.0\mu m$
length of the output guide	$50.0\mu m$
taperlength	$1000.0\mu m$
thickness of the guiding layers	$0.014\mu m$

4.4. THE ASYMMETRICAL TAPER STRUCTURE

The spatial refractive index geometry of the 1D-model of the taper to fabricate is given in Figure 25 and its data are assembled in Table IX. Figure 26 shows the simulated field distribution using $TOL = 0.01$ and Figure 27 displays the appropriate distribution of nodes. It is seen that the structure behaves as expected, only a small amount of power is radiated away at the butt coupling to the output waveguide (note that the amplitude is logarithmically scaled). This impression is backed quantitatively by the computed power coupling coefficient given in Figure 28. At the output of the structure we have a power coupling efficiency of about 0.96.

Table X describes the numerical effort using different tolerances and the obtained coupling coefficients.

It is found again that a $TOL = 0.08$ is sufficient for practical purposes, if

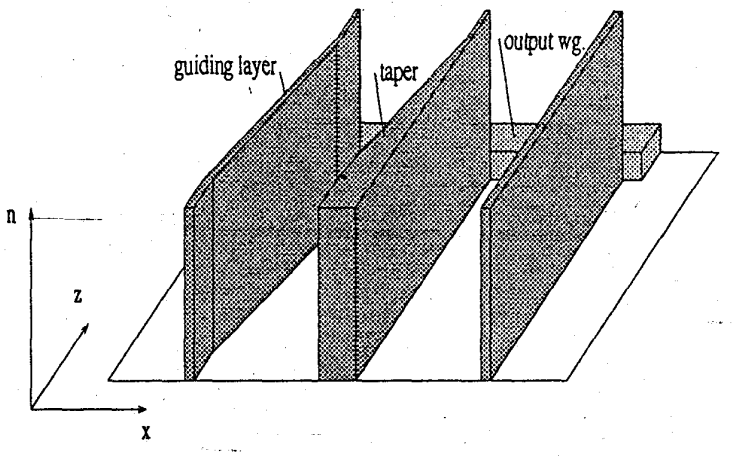


FIG. 25. Complete 1D model of the taper coupling

TABLE X. Asymmetrical taper: Numerical effort and results using different tolerances

TOL	CPU/s	steps	η
0.08	25.0	110	0.9655
0.06	33.9	124	0.9690
0.04	44.1	152	0.9653
0.02	99.7	239	0.9667
0.01	216.0	364	0.9634

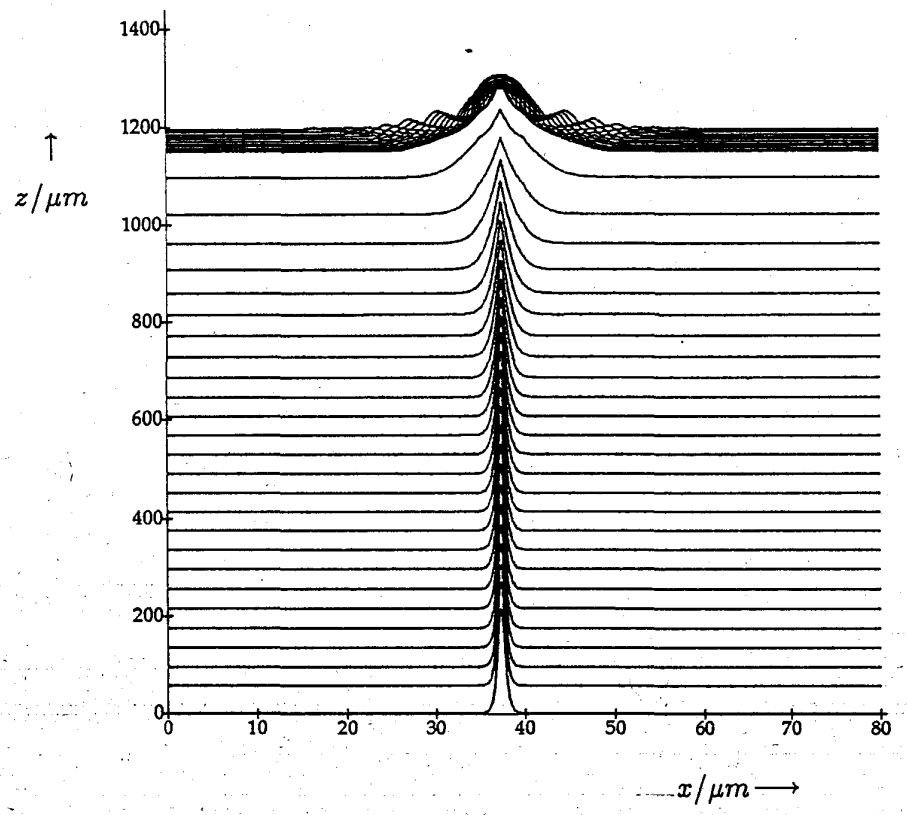


FIG. 26. Asymmetrical Taper: fielddistribution, $TOL = 0.01$, amplitude logarithmic scaled

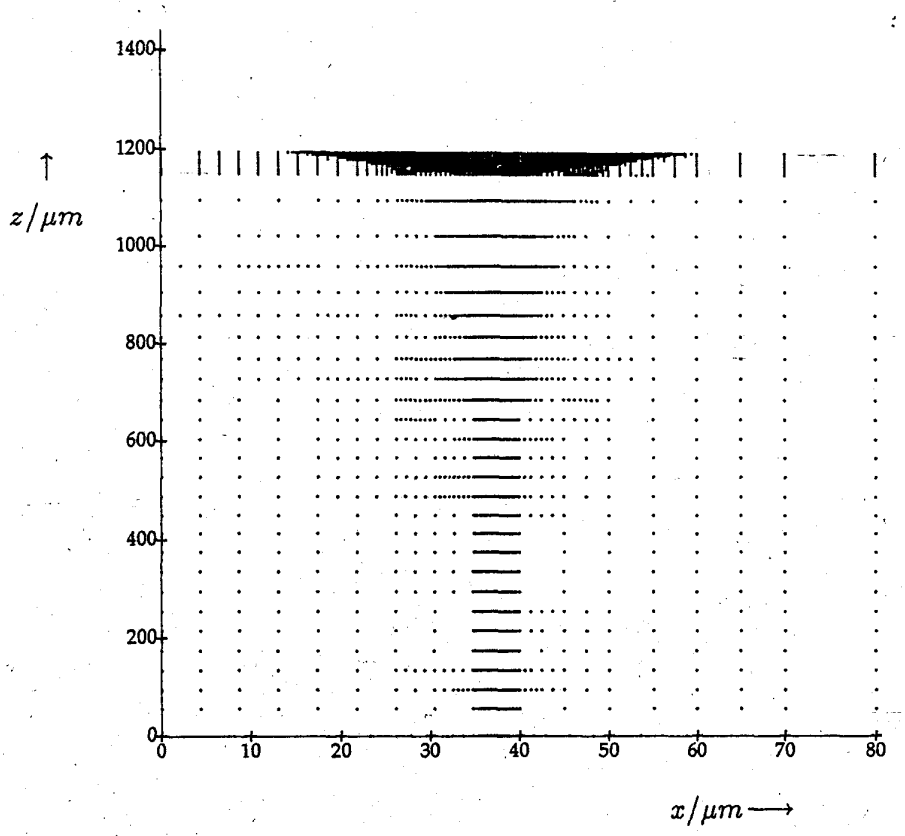


FIG. 27. Asymmetrical taper: distribution of nodes, $TOL = 0.01$

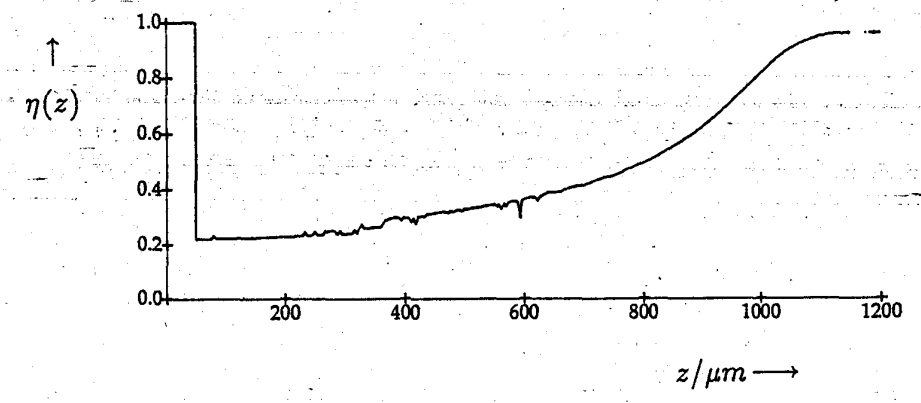


FIG. 28. Asymmetrical taper: Power coupling coefficient

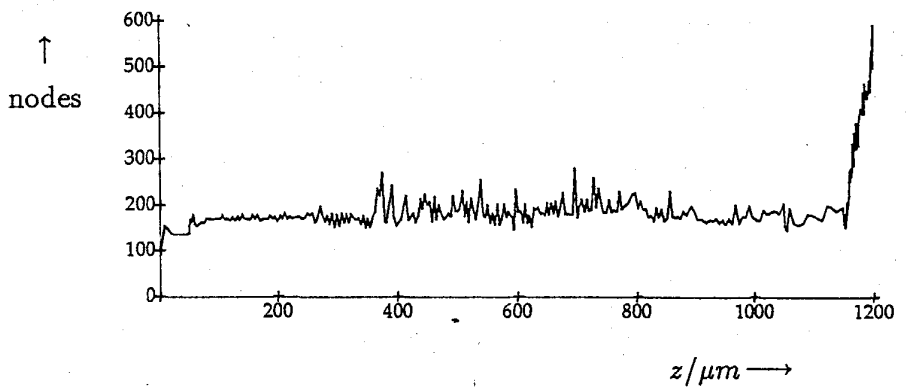


FIG. 29. Asymmetrical taper: number of nodes

only the power coupling coefficient is considered. To complete the description of the adaptive algorithm, the evolution of the number of nodes and of the stepsize are displayed in the Figures 29 and 30, respectively.

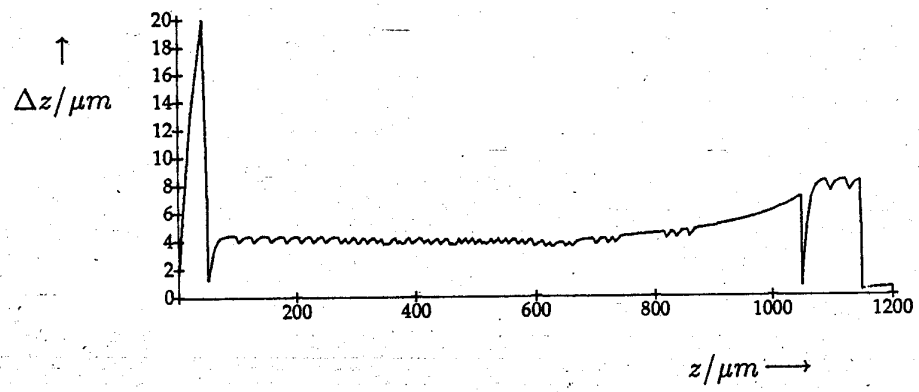


FIG. 30. Asymmetrical taper: evolution of Δz. The abrupt Δz-changes indicate abrupt geometrical changes (comp. Figure 25).

Conclusions

The following main results have been presented:

1. A self-adaptive finite element method to simulate the propagation of light in integrated optics devices was implemented following Bornemann's adaptive Rothe method.

2. It was shown that the the implicit midpoint discretization in the longitudinal direction supplied the desired behavior of the power evolution.

3. The observation of the deviation of the power from a constant value can be used to monitor the validity of Fresnel's approximation.

The next steps will be to integrate the Helmholtz equation rather than Fresnel's approximation and to extend the algorithm to 2D cross sections.

Acknowledgment

This work has been performed within a cooperation between the Heinrich-Hertz-Institut Berlin (HHI) and the Konrad-Zuse-Zentrum Berlin (ZIB) and has been supported by the Deutsche Bundespost Telekom under contract 'Waveguide Tapers for OEIC based on InP'.
The author gratefully acknowledges the guidance and encouragement of Prof. P. Deuflhard and Dr. F. A. Bornemann (ZIB) and of Dr. H.-P. Nolting (HHI).

REFERENCES

[1] Babŭska, I. , Rheinboldt, W. C.: *Error Estimates for Adaptive Finite Element Computations.* SIAM J. Numer. Anal. 15, 736-754 (1979)

[2] Bornemann, F. A. : *An Adaptive Multilevel Approach to Parabolic Equations* 1. *General Theory and Implementation.* IMPACT Comput. Sci. Engrg. 2, 279-317 (1990).

[3] Bornemann, F. A. : *An Adaptive Multilevel Approach to Parabolic Equations* 2. *Variable–Order Time Discretization Based on a Multilevel Error Correction.* IMPACT Comput. Sci. Engrg. 3, 93-122 (1991).

[4] Deuflhard, P. : *Order and Stepsize Control in Extrapolation Methods.* Numer. Math. 41, 399-422 (1983).

[5] Deuflhard, P. : *Recent Progress in Extrapolation Methods for Ordinary Differential Equations.* SIAM Rev. 27, 505-536 (1985).

[6] Feit, M. D. , Fleck, J. A. : *Light propagation in graded-index optical fibers.* Appl. Optics 17, 3390-3998 (1978).

[7] Gerdes, J. , Pregla, R. : *New beam-propagation algorithm based on the method of lines.* Techn. Digest on Int. Photonics Research, 5, 29-30 (1990)

[8] Marcuse, D.: *Theory of Dielectric Optical Waveguides.* Academic Press, New York (1974).

[9] Nolting, H. -P. : *Theoretical Investigations of Optical Tapers on InGaAsP/InP.* European Cooperation in the Field of Scientific and Technical Research, Action COST 240 'Modeling and Measuring Advanced Photonic Telecommunication Components', 1st workshop Zurich, 11th Nov. 1991.

[10] Koch, T. B. , März, R. , and Davies, J. B. : *Beam propagation method using z-transient variational principle.* in Proc 16th European Conf. Opt. Commun. (ECOC), 1 Amsterdam, The Netherlands, 163-166 (1990)

[11] Ratowsky, R. , Fleck, J. : *Accurate numerical solution of the Helmholtz equation by iterative Lanczos reduction.* Optics Letters, 11, 787-789 (1991).

[12] Simonyi, K. : *Theoretische Elektrotechnik.* VEB Deutscher Verlag der Wissenschaften, Berlin (1977).

[13] Splett, A. : Private Communication.

[14] Strang, G. , Fix, J. G. : *An Analysis of the Finite Element Method.* Prentice-Hall, Inc. , Englewood Cliffs (1973).

[15] Yevick, D. ,Hermansson, B. : *Efficient beam propagation techniques.* IEEE J. Quant. Elec. , 26, Vol. 26, No. 1, 109-112, (1990).

[16] Zienkiewicz, O. C. : *The Finite Element Method.* 3rd edition, McGraw-Hill, (1982).