Efficient beam-propagation method based on Padé approximants in the propagation direction

Yih-Peng Chiou

Department of Electrical Engineering, National Taiwan University, Taipei 106-17, Taiwan

Hung-chun Chang

Department of Electrical Engineering and Graduate Institute of Electro-Optical Engineering, National Taiwan University, Taipei 106-17, Taiwan

Received February 25, 1997

We propose a novel beam-propagation method (BPM) based on the higher-order Padé approximants in both the transverse and the longitudinal directions. Extending the Padé approximation to the propagation direction and adopting the multistep method increase the programming effort by only a small amount. Compared with the conventional BPM’s, this method is more accurate and efficient. The accuracy and the propagation step size are well predicted by a phase analysis. © 1997 Optical Society of America

The beam-propagation method (BPM) first derived from the paraxial approximation of the scalar wave equation has been widely used in the investigation of optoelectronic devices. The original BPM was based on the fast-Fourier-transform technique and was restricted to paraxial, scalar, and low-index contrast cases. Later, BPM’s based on the finite-difference technique were proposed. The finite-difference BPM’s are more efficient and can be applied to structures with larger contrast in the refractive index. Furthermore, semivectorial and full-vectorial BPM’s were developed, and wide-angle (WA) BPM’s were proposed to get rid of the paraxial restriction. However, the step size in the propagation direction must be small enough, especially in the WA BPM’s because of the rapid variation of waves at different propagation directions, to ensure the replacement of the partial differentiation by the finite difference.

The conventional finite-difference BPM’s are based on explicit or implicit formulas, which are basically Padé approximants of the first two lower orders in the propagation direction. Extending the Padé approximation to the propagation direction and using the Padé approximants in both the transverse and the longitudinal directions. The step size can be much larger, and the accuracy is much higher.

From Maxwell’s equations and neglecting the refractive-index variation in the propagation direction, we obtain the following wave equation in terms of the transverse fields:

\[
\frac{\partial^2 E_t}{\partial z^2} + \nabla^2 E_t + k_0^2 n^2 E_t + \nabla \left( \nabla \frac{n^2}{n} E_t \right) = 0,
\]

where \( k_0 \) is the wave number in free space, \( n \) is the refractive-index distribution, and the subscript \( t \) denotes the transverse direction. Considering propagating waves and discretizing the transverse direction, we have

\[
\frac{d^2 \Phi}{dz^2} + [L] \Phi = 0,
\]

where \([L]\) is the characteristic matrix and \( \Phi \) represents a column vector composed of nodal electrical fields after discretization. From Eq. (2) and considering the forward-propagating field, we have

\[
\Phi(z_0 + \Delta z) = \exp(-j\sqrt{[L]} \Delta z) \Phi(z_0).
\]

The function of \([L]\) is solved by

\[
\exp(-j\sqrt{[L]} \Delta z) = [V]^{-1} \exp(-j\sqrt{[\Lambda]} \Delta z) [V]
\]

if \([L] = [V]^{-1}[\Lambda][V]\), where \([\Lambda]\) is a diagonal matrix whose elements are the eigenvalues of \([L]\) and where the columns of \([V]\) are the corresponding eigenvectors. These eigenvalues and eigenvectors of the characteristic matrix are in fact the propagation constants and modes of the structure. However, solving an eigen-system is an inefficient procedure that requires much computation time and memory, especially for non-Hermitian characteristic matrices.

To avoid the need to solve the eigenvalue problem, efforts to approximate the square root of the characteristic matrix, the so-called WA BPM’s, have been made. One approach to solution of the problem is to use the Padé approximants with the original functions replaced by the rational functions.

Assuming that \( \Phi(z) = \exp(-jk_0 \hat{n} z) \phi(z) \) and substituting this equation into Eq. (2), we have the recurrence relation

\[
\frac{d}{dz} = \left[ L - k_0^2 \hat{n} \right] \frac{d^2 \phi}{dz^2} = [P],
\]

where \( \hat{n} \) is the reference refractive index and \([P]\) is a rational function of \([L]\). Thus the propagation equation becomes

\[
\frac{d \phi}{dz} = [P] \phi(z).
\]

Note that the Padé approximants with \([P]\) treat the field variation in the transverse direction. Conventionally, the \( z \) direction is discretized directly in Eq. (6)
to yield
\[ \phi(z + \Delta z) = \frac{2[D] - j\Delta z[N]}{[D] + j\Delta z[N]} \phi(z), \]
where \([N] \) and \([D] \) are polynomials of \([L] \), and \([P] = -j[N]/[D] \). Alternatively, from Eq. (6) we have
\[ \phi(z + \Delta z) = \exp((P)\Delta z)\phi(z) = \exp(\eta)\phi(z). \]

Here we propose the use of Padé approximants to approximate the exponential function instead of taking the finite difference in Eq. (6) directly. The Padé approximants of \(\exp(\eta)\) of lower orders are given by
\[ P_{22}(\eta) = 1 + \eta, \]
which is the commonly used explicit scheme,
\[ P_{211}(\eta) = (2 + \eta)/(2 - \eta), \]
which is the frequently used Crank–Nicholson implicit scheme, and
\[ P_{22}(\eta) = (12 + 6\eta + \eta^2)/(12 - 6\eta + \eta^2), \]
\[ P_{233}(\eta) = (120 + 60\eta + 12\eta^2 + \eta^3)/
(120 - 60\eta + 12\eta^2 - \eta^3). \]
The truncation error for the replacement of \(\exp(\eta)\) by \(P_{pq}(\eta)\) is
\[ \exp(\eta) = P_{pq}(\eta) + O(\eta^{p+q+1}). \]
From Eqs. (8)–(10) and (13) the truncation errors for the explicit and the Crank–Nicholson implicit schemes are \(O(\Delta z^2)\) and \(O(\Delta z^3)\), respectively. Consequently, if the higher-order Padé approximants are used, more-accurate results can be obtained and the propagation step \(\Delta z\) can be much larger. In this Letter we use \(P_{22}\) and \(P_{233}\), whose truncation errors are \(O(\Delta z^5)\) and \(O(\Delta z^7)\), respectively. The multistep method can be adopted here to make use of the sparsity of the characteristic matrix. From Eqs. (8)–(13) we have
\[ \phi(z + \Delta z) = \prod_{r=0}^{p}\frac{1 + c_r[L]}{1 + c_r[L]} \phi(z), \]
\[ \phi^{(p+1)pr} = \left(\frac{1 + c_p[L]}{1 + c_p[L]}\right) \phi^{pr}, \]
where \(p\) is the order of the Padé approximant of the exponential function and \(r\) is the order of the Padé approximant of the square root. We verified that the proposed method is unconditionally stable by using the von Neuman stability analysis on the unitary operator in Eq. (15).

We define the normalized step as the phase in Eq. (8):
\[ \zeta = ([N]/[D])\Delta z. \]
In the paraxial regime the normalized step is \(\zeta = (\beta^2 - k_0n^2)\Delta z/2k_0n\), and in the WA cases \(\zeta = (\beta - k_0n)\Delta z\), where \(\beta\) is the propagation constant of the propagating wave. Thus, if the difference between \(\beta\) and \(k_0n\) is larger, the propagation step size \(\Delta z\) should be smaller to yield the same accuracy, especially in the WA cases because of the much larger index difference.

Figure 1 shows the relative phase errors for various Padé approximants \(P_2\) in the propagation direction. The relative phase error is defined as the phase error after propagation through one unit step and is given by
\[ \epsilon(\zeta) = \left|\frac{\phi_{\text{calculated}} - \zeta}{\zeta}\right|, \]
where \(\phi_{\text{calculated}}\) is the phase calculated by the various Padé approximants. The open circles show the results with \(P_{10}\), corresponding to an implicit scheme in Eq. (9). Solid, dotted, and dashed curves are results with \(P_{11}\), \(P_{22}\), and \(P_{33}\) in Eqs. (10), (11), and (12), respectively. The relative phase error gets smaller when the step size or \(\zeta\) decreases for all orders, as we expect. The accuracy of \(P_2\) of higher orders is much greater for the same normalized phase. For example, with the same phase error of \(10^{-4}\), the step sizes of \(P_{22}\) and \(P_{33}\) are \(\sim 10\) and \(\sim 100\) times larger, respectively, than that of \(P_{22}\), whereas the corresponding computation times are only doubled and tripled.

Because the propagation phenomenon of optical waves in a directional coupler is well known and can be obtained exactly, it is used in what follows to demonstrate the proposed method. The structure is shown as an inset in Fig. 2. The wavelength \(\lambda\) is taken to be \(1\mu\text{m}\), the core width \(W\) is \(2\mu\text{m}\), the separation \(S\) between the two cores is \(1\mu\text{m}\), and the refractive indices of the core and the cladding are \(n_1 = 1.55\) and \(n_0 = 1.50\), respectively. The input field is the fundamental mode in waveguide \(a\), whose effective index \(n_e = \beta/k_0 = 1.540008\).

Figure 2 shows the power variation in waveguide \(a\), which we calculated by the paraxial approximation, using \(P_{22}\) and \(P_{222}\) after \(15,000–16,000\mu\text{m}\) propagation. The reference index \(n\) is taken to be the effective index \(n_e\) of a single waveguide, which satisfies the paraxial assumption. The solid curve is calculated by \(P_{222}\) with \(\Delta z = 1\mu\text{m}\), which is treated as the convergent limit. The dotted and dashed curves are calculated by \(P_{221}\) with \(\Delta z = 20\) and \(40\mu\text{m}\), respectively. The boxes are calculated by \(P_{222}\) with \(\Delta z = 200\mu\text{m}\). The normalized steps \(\zeta\) are \(0.007, 0.14, 0.28,\) and \(1.4\) for \(\Delta z = 1, 20, 40,\) and \(200\mu\text{m}\), respectively. As \(\epsilon(\zeta)\) for \(P_{222}\) at \(\zeta = 1.4\) is approximately equal to that for \(P_{221}\) at \(\zeta = 0.14\) (Fig. 1), the two results would have
approximate accuracy. Figure 2 indeed shows that with the same accuracy the step size with $Pz_{22}$ can be approximately ten times larger than that with $Pz_{11}$. Our calculation also shows that using $Pz_{33}$ can increase the step size one more order than using $Pz_{22}$ with the same accuracy. The huge step in this example results from the fact that $\bar{n}$ is very close to $n_e$ of the coupler. Our calculations show that if $\bar{n}$ is slightly different from $n_e$, for example, $\bar{n} = 0.98n_e$, the result will deviate significantly from the exact result, even for tiny steps.

Figure 3 shows the power variation in waveguide $a$ in WA cases, which is calculated with $Pz_{11}$ and $Pz_{33}$ after 15,000–16,000-µm propagation. The square root of the characteristic matrix is approximated with (3,3) Padé. The solid curve is calculated by use of $Pz_{33}$ with $\Delta z = 1$ µm and $\bar{n} = n_e$, which is taken as the convergent limit whose normalized step $\zeta$ is 0.007. For the dotted and dashed curves, $\bar{n}$ is taken to be $0.7n_e$, and their normalized steps $\zeta$ are 0.145 and 1.45 for $\Delta z = 0.05$ and 0.5 µm, respectively. Because the relative phase error for $Pz_{11}$ with $\Delta z = 0.05$ µm is ten times larger than that for $Pz_{33}$ with $\Delta z = 0.5$ µm (Fig. 1), the result calculated by $Pz_{11}$ is much worse than that by $Pz_{33}$. To achieve the same accuracy as with $Pz_{33}$ requires a step size for $Pz_{11}$ of $0.0014$ µm (Fig. 1), as is validated by our computation. Thus the step size for $Pz_{33}$ is ~35 times larger than that for $Pz_{11}$ at the same accuracy; the computation time is only tripled.

In conclusion, we have proposed an efficient and accurate beam-propagation method based on Padé approximants of higher orders in the propagation direction. In this method the step size can be much larger and the accuracy increased compared with those for conventional BPM’s. Both paraxial and wide-angle cases were demonstrated. The relation between the numerical accuracy and the propagation step size used is well predicted by that between the relative phase error and the normalized step.

This study was supported by the National Science Council of the Republic of China under grant NSC86-2215-E-002-008.

References