10

Propagation equations for Finite-Difference Methods

Previous chapters introduced several methods that can serve as components in more complex beam propagation implementations. For some approaches, namely the spectral ones, it was not even necessary to speak about propagation equations, because the superposition principle was sufficient to construct the method from scratch. For other cases, such as various versions of the Crank-Nicolson method or the Alternating Direction Implicit method, we chose a simplified geometry, mostly without waveguiding structure that is characteristic of the BPM in general. Now it is the time to see how these tools can be applied in waveguiding structures with non-trivial geometry. To this end, we re-visit derivation of the underlying propagation equations, and have a closer look on how they can be discretized. Then we discuss semi-vectorial and full-vector versions of BPM, together with description of implementation of ADI and iterated Crank-Nicolson methods. The main goal of this chapter is to demonstrate how the algorithms discussed so far make it possible to build more complex BPM simulators. For simplicity, the following treatment is restricted to isotropic and non-magnetic media.

10.1 Electric field equations

We assume a single-frequency, or monochromatic regime, and transform the Maxwell equations into frequency representation:

$$\nabla \times \mathbf{E} = -i\omega\mu_0\mathbf{H} \tag{10.1}$$

$$\nabla \times \mathbf{H} = +i\omega n^2\epsilon_0\mathbf{E} \tag{10.2}$$

Here, n represents the refractive index of the structure to be simulated, and is implicitly assumed to depend on the location; it may vary as a continuous function of spatial coordinates, and it can exhibit sharp interfaces between domains of different materials. The angular frequency ω serves as a fixed parameter in this system, and the refractive index should take values specific for this frequency. This system of equations is our stating point, and we take the usual route forward, namely through the vectorial Helmholtz equations for both electric and magnetic fields. Let us start with the electric equations first.

Application of the *curl* operator to the left- and right-hand sides of the first equation, while eliminating $\nabla \times \mathbf{H}$ with the help of the second we obtain

$$\nabla \times \nabla \times \mathbf{E} - n^2k^2\mathbf{E} = 0 \tag{10.3}$$

Next, the well-known differential identity

$$\nabla \times \nabla \times = -\Delta + \nabla\nabla\cdot \tag{10.4}$$

used in the previous equation leads us to

$$\Delta\mathbf{E} + n^2k^2\mathbf{E} = \nabla\nabla\cdot\mathbf{E}\ . \tag{10.5}$$

It is often convenient to treat transverse components of the field (i.e. the x, y vectorial components) and the longitudinal component E_z separately. This will help to reduce the system to $\mathbf{E}_x, \mathbf{E}_y$ alone while keeping $\mathbf{E}_z$ implicitly accounted for. So we split both nabla operators on the right-hand side into longitudinal and transverse components, and take the transverse part of the whole equation

$$\Delta\mathbf{E}_t + n^2k^2\mathbf{E}_t = \nabla_t(\nabla_t\cdot\mathbf{E}_t + \partial_z E_z)\ , \tag{10.6}$$

where vectors carrying subscript t are understood to live in the subspace spanned by x and y. Similarly, the longitudinal part of the equation becomes

$$\Delta\mathbf{E}_z + n^2k^2\mathbf{E}_z = \nabla_z(\nabla_t\cdot\mathbf{E}_t + \partial_z E_z)\ . \tag{10.7}$$

Next, one eliminates the $\mathbf{E}_z$ with the help of the divergence equations. Note that this is exactly the point at which the divergence equations, or equivalently the initial constraints, are embedded into the BPM machinery. Without free charges floating around, the electric displacement divergence must vanish, so

$$\nabla\cdot\mathbf{D} = \epsilon_0\nabla\cdot(n^2\mathbf{E}) = 0 \tag{10.8}$$

which explicitly means that

$$\nabla_t\cdot(n^2\mathbf{E}_t) + \partial_z n^2 E_z + n^2\partial_z E_z = 0\ . \tag{10.9}$$

Here we have partial z derivative acting on both the refractive index function and on the electric z-component of the field. Because only $\partial_z\mathbf{E}_z$ appears in the equation before, one would like to remove the second term. Fortunately, this is an excellent approximation most of the time; While the last term can be estimated as $k_0\mathbf{E}_z = \omega/c\mathbf{E}_z = 2\pi/\lambda\mathbf{E}_z$, the second varies as $1/\ell\mathbf{E}_z$ with ℓ standing for the characteristic length in direction of beam propagation over which the refractive index changes significantly. If the latter is much longer that the wavelength, the longitudinal field derivative can be isolated as

$$\partial_z\mathbf{E}_z \approx -\frac{1}{n^2}\nabla_t\cdot(n^2\mathbf{E}_t) \tag{10.10}$$

This is an expression which will be returned into Equations (10.6,10.7). Note that this is exact if n does not change along propagation direction, which is in fact very often the geometry encountered in the BPM context. This approximation (or an exact relation) leads to a wave equation equation for transverse electric-field components,

$$\Delta\mathbf{E}_t + n^2k^2\mathbf{E}_t = \nabla_t[\nabla_t\cdot\mathbf{E}_t - \frac{1}{n^2}\nabla_t\cdot(n^2\mathbf{E}_t)]\ , \tag{10.11}$$

and also for the longitudinal electric field,

$$\Delta\mathbf{E}_z + n^2k^2\mathbf{E}_z = \nabla_z[\nabla_t\cdot\mathbf{E}_t - \frac{1}{n^2}\nabla_t\cdot(n^2\mathbf{E}_t)]\ . \tag{10.12}$$

The following, explicit component form emphasizes the symmetry of the system:

$$\nabla^2E_x + \frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2E_x)\right] + \frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2E_y)\right] - \frac{\partial^2E_x}{\partial x^2} - \frac{\partial^2E_y}{\partial x\partial y} + k_0^2n^2E_x = 0 \tag{10.13}$$

$$\nabla^2 E_y + \frac{\partial}{\partial y}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2 E_y)\right] + \frac{\partial}{\partial y}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2 E_x)\right] - \frac{\partial^2 E_y}{\partial y^2} - \frac{\partial^2 E_x}{\partial y \partial x} + k_0^2 n^2 E_y = 0 \qquad (10.14)$$

$$\nabla^2 E_z + \frac{\partial}{\partial z}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2 E_x)\right] + \frac{\partial}{\partial z}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2 E_y)\right] - \frac{\partial^2 E_x}{\partial x \partial z} - \frac{\partial^2 E_y}{\partial y \partial z} + k_0^2 n^2 E_z = 0 \qquad (10.15)$$

This is a point of departure for various discrete grid representations of beam propagation equations.

10.2 One-way propagation: A single envelope approximation

An important point to note at this stage is that the above wave equation can not be directly used as a propagation equation for a beam, despite that the longitudinal field component has been eliminated. This is because the equation is of second order with respect to the desired propagation direction z. One might be first tempted to say that a second order system can be equivalently formulated as a richer (i.e. having more unknown functions) first-order system. This would however quickly reveal an obstacle. One would need two initial conditions to solve a second-order equation, and that is not available when only a beam profile is specified at the entrance into a device which is to be simulated. Estimating the derivative initial condition would be also utterly useless, because the numerics would quickly show that the equation supports both forward and backward propagating waves, and would "create" an unwanted backward solution from numerical noise. An alternative way to look at this is to realize that instead of an initial value problem, it is a boundary value problem that is naturally connected with the Helmholtz equation; What is specified in practice is the input beam on the entrance of a device and *absence* of an incoming beam on the output from the device. A shooting method can be in principle used to solve it by repeating (shooting) solutions staring from the input, and correcting an unknown initial derivative condition until a zero incoming beam is obtained at the output. However, beam propagation equation systems are typically large and this would be very difficult. Thus we see that to proceed toward a first-order evolution system of equations to describe beams, we need to acknowledge the fact that the wave equation describes waves propagating in "all" directions, and remove the undesired backward propagating field.

There are several way to deal with this problem, and accordingly various versions of beam propagation methods. We will discuss three different treatments resulting in three subsets of BP methods, including so called wide-angle BPM. In the following section we start with the simplest approach which will produce numerical methods suitable for devices in which beams propagate at no too large angles, and remain nearly paraxial.

The basic idea is to remove from the computational picture the fast changes related to the carrier wave. At the same time, an advantage is taken of the fact that it is the carrier wave that encodes the direction of propagation into a wavepacket. Assuming a general electric field of a waveform propagating in a predominant direction (let it be z as usual), it can be represented as

$$\mathbf{E}_i(x, y, z, t) = \mathcal{E}_i(x, y, z) e^{ik_{ref} z - i\omega t} + c.c. \qquad (10.16)$$

where $\mathcal{E}_i(x, y, z)$ is the field envelope function. The exponential function stands for the spatial variation of the electric field intensity due to the carrier wave — this is the part we do not want to represent on a numerical grid since it varies significantly over length scales comparable to the wavelength. On the other hand, the envelope can be a slow function of its z-argument.

Whether it actually is or is not, depends on the choice of k_{ref} — if this really corresponds to the predominant wavenumber of the wavepacket, then the envelope must only take care of slower (with respect to z) variations. Correspondingly, a single step of the simulated BPM evolution may potentially be significantly longer than the light wavelength.

Here it should be emphasized that the concrete value of k_{ref} is *chosen* rather than given. In particular, k_{ref} did not appear and has no place in the Maxwell equations. Another important aspect to note is that the envelope function is complex-valued. However, by now we are used to this "complexification", and know that the apparent complication is out-weighted by the fact that a much coarser grid is usually sufficient to represent $\mathcal{E}_i(x, y, z)$ in comparison to that needed for $\mathbf{E}_i(x, y, z)$.

Now we arrive at the point where semi-vectorial and Wide-Angle BPM separate. Let us follow the simpler path for now, and return to WA-BPM when we know more about practicalities of discretization in the presence of interfaces and variable refractive index.

After inserting the envelope representation of the optical field into Helmholtz, we drop the second z-derivative of $\mathcal{E}_i$ — this is the well-know slowly evolving envelope approximation (SVEA). This significantly simplifies the whole system, because SVEA turns it into an initial value problem for the following first-order evolution equations

$$\frac{\partial \mathcal{E}_x}{\partial z} = \hat{A}_{xx}\mathcal{E}_x + \hat{A}_{xy}\mathcal{E}_y \tag{10.17}$$

$$\frac{\partial \mathcal{E}_y}{\partial z} = \hat{A}_{yy}\mathcal{E}_y + \hat{A}_{yx}\mathcal{E}_x \tag{10.18}$$

where the short-hands on the right represent these operators

$$\hat{A}_{xx}\mathcal{E}_x = \frac{i}{2k_{ref}}\left(\frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2\mathcal{E}_x)\right] + \frac{\partial^2 \mathcal{E}_x}{\partial y^2} + k_0^2(n^2 - n_{ref}^2)\mathcal{E}_x\right) \tag{10.19}$$

$$\hat{A}_{yy}\mathcal{E}_y = \frac{i}{2k_{ref}}\left(\frac{\partial}{\partial y}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2\mathcal{E}_y)\right] + \frac{\partial^2 \mathcal{E}_y}{\partial x^2} + k_0^2(n^2 - n_{ref}^2)\mathcal{E}_y\right) \tag{10.20}$$

$$\hat{A}_{xy}\mathcal{E}_y = \frac{i}{2k_{ref}}\left(\frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2\mathcal{E}_y)\right] - \frac{\partial^2 \mathcal{E}_y}{\partial x \partial y}\right) \tag{10.21}$$

$$\hat{A}_{yx}\mathcal{E}_x = \frac{i}{2k_{ref}}\left(\frac{\partial}{\partial y}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2\mathcal{E}_x)\right] - \frac{\partial^2 \mathcal{E}_x}{\partial y \partial x}\right) \ . \tag{10.22}$$

This is a closed system for the transverse field components. Also note that only refractive index variations contribute to mixing between the two polarizations. If there are material interfaces within the computational domain, some quantities may not exist directly *at* an interface. However, the fields that appear in this formulations are actually continuous across the material interfaces. For example, consider a sharp boundary between two materials that is perpendicular to x. Then $n^2\mathcal{E}_x$ that appears in A_{xx} represents the normal component of the electric induction which we know is continuous (in the absence of surface charges, of course). So is the normal derivative as implied by the divergence equation (in the absence of free charges). These properties make the discretization easier.

The evolution system for the longitudinal component is somewhat different, and reads as

$$\frac{\partial \mathcal{E}_z}{\partial z} = \hat{A}_{zz}\mathcal{E}_z + \hat{B}(\mathcal{E}_x, \mathcal{E}_y) \tag{10.23}$$

with

$$\hat{A}_{zz}\mathcal{E}_z = \frac{i}{2k_{ref}}\left(\frac{\partial^2\mathcal{E}_z}{\partial x^2} + \frac{\partial^2\mathcal{E}_z}{\partial y^2} + k_0^2(n^2 - n_{ref}^2)\mathcal{E}_z\right) \tag{10.24}$$

$$\hat{B}(\mathcal{E}_x, \mathcal{E}_y) = \frac{i}{2k_{ref}}\left(\frac{\partial}{\partial z} + ik_{ref}\right)\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2\mathcal{E}_x) + \frac{1}{n^2}\frac{\partial}{\partial y}(n^2\mathcal{E}_y) - \frac{\partial\mathcal{E}_x}{\partial x} - \frac{\partial\mathcal{E}_y}{\partial y}\right] \tag{10.25}$$

Here we can see that to solve for the longitudinal component, we must know the transverse components first. These then drive, or control evolution of $\mathcal{E}_z$. However, most of the time, only the transverse components are simulated. Besides the fact that $\mathcal{E}_z$ can be eliminated, an additional reason for this is that we have already committed to the SVEA, which in turn "guarantees" that the longitudinal component is small.

10.3 Discretization of BPM equations

The discretization procedure of the equations derived above is described in detail next. In its entirety it would be rather long and repetitive, because there are only a couple of "patterns" that are actually needed. This is what we concentrate on now.

First, consider typical term from the evolution system, e.g.

$$\hat{A}_{xx}\mathcal{E}_x = \frac{i}{2k_{ref}}\frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2\mathcal{E}_x)\right] + \ldots \tag{10.26}$$

Our next task is to design its finite-difference approximation, which is unfortunately not unique. It depends on how smooth the refractive index function is considered to be. We will assume that sharp interfaces may be present in the structure, and that we can not rely on smooth approximations for the permittivity.

Our guiding principle is what we have learned with the Crank-Nicolson and Maxwell discretization; it is usually best to keep finite-difference expression as symmetric as possible. So, for the discrete grid point point i, we can pretend for now that necessary quantities are known for locations in-between the spatial grid points, and write the derivative as follows:

$$\hat{A}_{xx}\mathcal{E}_x \approx \frac{1}{\Delta x}\left[\left(\frac{1}{\epsilon}\partial_x(\epsilon E_x)\right)\Big|_{i+1/2} - \left(\frac{1}{\epsilon}\partial_x(\epsilon E_x)\right)\Big|_{i-1/2}\right], \tag{10.27}$$

which expression is of course second order accurate in Δx. Immediately, we face the question of how to approximate $1/\epsilon$. Remember that it is usually only specified at grid-point locations, and an analytic formula may not be available. In other words, $1/\epsilon|_{i+1/2}$ must be written in terms of the permittivity samples at grid points. Let us therefore take

$$\left(\frac{1}{\epsilon}\right)\Big|_{i+1/2} \approx \frac{1}{2\epsilon_i} + \frac{1}{2\epsilon_{i+1}}\ . \tag{10.28}$$

This is of course nothing but a linear interpolation for the function $1/\epsilon$. Next, using standard symmetric first-order derivative approximation we get

$$\hat{A}_{xx}\mathcal{E}_x \approx \frac{1}{\Delta x}\left[\left(\frac{1}{2\epsilon_i} + \frac{1}{2\epsilon_{i+1}}\right)\frac{(\epsilon E_x)_{i+1} - (\epsilon E_x)_i}{\Delta x} - \left(\frac{1}{2\epsilon_i} + \frac{1}{2\epsilon_{i-1}}\right)\frac{(\epsilon E_x)_i - (\epsilon E_x)_{i-1}}{\Delta x}\right] \tag{10.29}$$

which simplifies to:

$$\hat{A}_{xx}\mathcal{E}_x \approx \frac{1}{\Delta x^2}\left[\frac{\epsilon_i+\epsilon_{i+1}}{2\epsilon_i}E_{i+1} - E_i\left(\frac{\epsilon_i+\epsilon_{i+1}}{2\epsilon_{i+1}} + \frac{\epsilon_i+\epsilon_{i-1}}{2\epsilon_{i-1}}\right) + \frac{\epsilon_i+\epsilon_{i-1}}{2\epsilon_i}E_{i-1}\right] . \tag{10.30}$$

As a quick sanity check, we can see that if the permittivity is constant, this expression reduces to the standard three-point approximation for $\partial_{xx}E$ as it should. Also note that this formula can be viewed and implemented as the standard Laplacian approximation with the coefficients $(1,-2,1)$ replaced by location-dependent factors constructed from the local permittivity $\epsilon = n^2$.

The other expressions in A_{xx} and A_{yy} can be discretized in the same way or they are familiar second derivatives we have encountered many times. There is however a different kind of term in the "non-diagonal" operator A_{xy}. This vanishes for constant permittivity and that suggests that we should discretize the whole of A_{xy} (and A_{yx}) in one go, keeping all resulting terms together such they can cancel in regions of constant ϵ. To do this, consider a grid point (i,j) and its four neighbors and add its four next-nearest neighbors on a square lattice. These points will contribute to the discretized expression.

Let us look first at the simpler part of A_{xy}

$$\frac{\partial^2}{\partial x\partial y}E = \frac{\partial}{\partial x}\frac{\partial}{\partial y}E .$$

Centering all finite difference expressions, we can write the approximation as follows:

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y}E \approx \frac{1}{2\Delta x}\left[\left(\frac{\partial}{\partial y}E\right)_{i+1,j} - \left(\frac{\partial}{\partial y}E\right)_{i-1,j}\right]$$

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y}E \approx \frac{1}{2\Delta x}\left[\frac{E_{i+1,j+1}-E_{i+1,j-1}}{2\Delta y} - \frac{E_{i-1,j+1}-E_{i-1,j-1}}{2\Delta y}\right] \tag{10.31}$$

This term is a combination of field values found at next-nearest neighbor grid points. The portion of A_{xy} that also depends on refractive index must cancel this in case of constant permittivity. The procedure is analogous, one only needs to keep track of the local permittivity. Let us show these steps explicitly for one of the terms, starting from

$$\frac{\partial}{\partial x}\left(\frac{1}{\epsilon}\frac{\partial \epsilon E}{\partial y}\right) \approx \frac{1}{2\Delta x}\left[\left(\frac{1}{\epsilon}\frac{\partial \epsilon E}{\partial y}\right)_{i+1,j} - \left(\frac{1}{\epsilon}\frac{\partial \epsilon E}{\partial y}\right)_{i-1,j}\right] ,$$

and centering the products around the left and right nearest neighbors we obtain

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y}E \approx \frac{1}{2\Delta x}\left[\frac{\epsilon_{i+1,j+1}E_{i+1,j+1}-\epsilon_{i+1,j-1}E_{i+1,j-1}}{2\epsilon_{i+1,j}\Delta y} - \frac{\epsilon_{i-1,j+1}E_{i-1,j+1}-\epsilon_{i-1,j-1}E_{i-1,j-1}}{2\epsilon_{i-1,j}\Delta y}\right] . \tag{10.32}$$

Putting (10.31) and (10.32) together gives the finite-difference expression of the non-diagonal operator

$$\begin{aligned} A_{xy}E_y \approx \frac{1}{4\Delta x\Delta y}\Bigg(&+\left[\frac{\epsilon_{i+1,j+1}}{\epsilon_{i+1,j}}-1\right]E_{i+1,j+1} - \left[\frac{\epsilon_{i-1,j+1}}{\epsilon_{i-1,j}}-1\right]E_{i-1,j+1} \\ &- \left[\frac{\epsilon_{i+1,j-1}}{\epsilon_{i+1,j}}-1\right]E_{i+1,j-1} + \left[\frac{\epsilon_{i-1,j-1}}{\epsilon_{i-1,j}}-1\right]E_{i-1,j-1}\Bigg) . \end{aligned} \tag{10.33}$$

Approximation for A_{yx} is of course completely analogous. For the sake of completeness, the corresponding formulas read as follows

$$\frac{\partial}{\partial y}\left(\frac{1}{\epsilon}\frac{\partial \epsilon E}{\partial x}\right) \approx \frac{1}{2\Delta y}\left[\left(\frac{1}{\epsilon}\frac{\partial \epsilon E}{\partial x}\right)_{i,j+1} - \left(\frac{1}{\epsilon}\frac{\partial \epsilon E}{\partial x}\right)_{i,j-1}\right] ,$$

$$A_{yx}E_x \approx \frac{1}{4\Delta x\Delta y}\left(+\left[\frac{\epsilon_{i+1,j+1}}{\epsilon_{i,j+1}}-1\right]E_{i+1,j+1}-\left[\frac{\epsilon_{i-1,j+1}}{\epsilon_{i,j+1}}-1\right]E_{i-1,j+1}\right.$$
$$\left.-\left[\frac{\epsilon_{i+1,j-1}}{\epsilon_{i,j-1}}-1\right]E_{i+1,j-1}+\left[\frac{\epsilon_{i-1,j-1}}{\epsilon_{i,j-1}}-1\right]E_{i-1,j-1}\right)\ . \quad (10.34)$$

Having discretized the differential expressions, one can write the finite-difference approximation for the whole right-and side of the evolution system for the two transverse field components.

These discrete approximations can be utilized in the Method of Lines without further modification. Another approach which does not require much further work is the iterated Crank-Nicolson method. Together with the alternating direction implicit method, they will be discussed in following sections.

10.4 Semi-Vectorial Approximation

Beam propagation problems are often solved in an approximation that neglects mixing between different polarization components of the optical field. This approximation consist in throwing away the nondiagonal operators A_{xy} and A_{yx} in

$$\frac{\partial\mathcal{E}_x}{\partial z}=\hat{A}_{xx}\mathcal{E}_x+\hat{A}_{xy}\mathcal{E}_y \quad (10.35)$$
$$\frac{\partial\mathcal{E}_y}{\partial z}=\hat{A}_{yy}\mathcal{E}_y+\hat{A}_{yx}\mathcal{E}_x \quad (10.36)$$

The explicit form of these equations was derived in Section 10.2. For example,

$$\hat{A}_{xy}\mathcal{E}_y=\frac{i}{2k_{ref}}\left(\frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2\mathcal{E}_y)\right]-\frac{\partial^2\mathcal{E}_y}{\partial x\partial y}+\right) \quad (10.37)$$

only contributes if the refractive index depends on y, since the above evaluates to

$$\hat{A}_{xy}\mathcal{E}_y=\frac{i}{2k_{ref}}\frac{\partial}{\partial x}\left[\frac{\partial\log n^2}{\partial y}\mathcal{E}_y\right]\ . \quad (10.38)$$

Similar term arises from the other non-diagonal (i.e. polarization scrambling) operator A_{yx}. These contributions are often small, and since without them the problem is significantly simplified, the approximation is invoked in many practical situations. It is called semi-vectorial, because it results in a a pair of equations for the transverse components,

$$\frac{\partial\mathcal{E}_x}{\partial z}=\hat{A}_{xx}\mathcal{E}_x \quad (10.39)$$
$$\frac{\partial\mathcal{E}_y}{\partial z}=\hat{A}_{yy}\mathcal{E}_y \quad (10.40)$$

which are de-coupled and can be therefore solved in isolation. This is akin to splitting a waveguide problem into TE and TM pair of sub-problems which are easier to solve. So what we have are essentially scalar equations, but together their solutions represent a good approximation of a vectorial field — hence the name semi-vectorial.

10.4.1 Implementation with ADI

Without the complication of the mixing terms, each of the above equations now can be solved straightforwardly with the ADI scheme. This requires fusion of the material we have covered in three of the previous sections, namely modification of the C-N approach, ADI formulation, and modification of the discretized Laplacian as discussed in this Chapter.

For the sake of completeness, let us bring these ingredients together into sort of a pseudo-code. It is sufficient to do this for the $\mathcal{E}_x$ polarization since the formulation is identical for $\mathcal{E}_x$. The explicit form of the evolution equation is

$$\frac{\partial \mathcal{E}_x}{\partial z} = \frac{i}{2k_{ref}} \left(\frac{\partial}{\partial x} \left[\frac{1}{n^2} \frac{\partial}{\partial x} (n^2 \mathcal{E}_x) \right] + \frac{\partial^2 \mathcal{E}_x}{\partial y^2} + k_0^2 (n^2 - n_r^2) \mathcal{E}_x \right) . \tag{10.41}$$

Let $E_{i,j}$ represent the discrete beam solution at the grid point $\{i, j\}$. Recall the discretization scheme derived for the "modified Laplacian" that reflects changing refractive index, and distinguish the corresponding discrete operator with tilde

$$(\tilde{\Delta}_{xx} E)_{i,j} = \frac{\epsilon_{i,j} + \epsilon_{i+1,j}}{2\epsilon_i} E_{i+1,j} - E_{i,j} \left(\frac{\epsilon_{i,j} + \epsilon_{i+1,j}}{2\epsilon_{i+1,j}} + \frac{\epsilon_{i,j} + \epsilon_{i-1,j}}{2\epsilon_{i-1,j}} \right) + \frac{\epsilon_{i,j} + \epsilon_{i-1,j}}{2\epsilon_{i,j}} E_{i-1,j} . \tag{10.42}$$

It must be remembered when implementing this in a program, that the coefficients of the Laplacian may depend on where along y (i.e. what value index j has) this is evaluated. It may be worthwhile to pre-calculate and store these coefficients for ease of use.
Derivative in the y direction stays standard,

$$(\Delta_{yy} E)_{i,j} = E_{i,j+1} - 2E_{i,j} + E_{i,j-1} . \tag{10.43}$$

Since there are indices that correspond to x and y directions in space, the notation is using comma to separate indices that belong to different spatial dimensions. Indices not separated by comma will represent matrix elements of an operator. For example,

$$(\tilde{\Delta}_{xx} E)_{i,j} = \sum_{k=i-1}^{i+1} (\tilde{\Delta}_{xx})_{ik,j} E_{k,j}$$

with the second index after comma being a reminder that the operator coefficient may depend on j. With these notation the first sub-step in an ADI will be

$$\frac{E^{n+1/2} - E^n}{\Delta z/2} = \frac{i}{2k_r \Delta x^2} \tilde{\Delta}_{xx} E^{n+1/2} + \frac{i}{2k_r \Delta y^2} \Delta_{yy} E^n + \frac{ik_0^2(\epsilon - n_r^2)}{4k_r} \left(E^{n+1/2} + E^n \right) . \tag{10.44}$$

Here, only $\tilde{\Delta}_{xx} E$ is taken implicit, and the diagonal term is averaged between steps as in the modified Crank-Nicolson.
The second step is completely analogous, this time it is $\Delta_{yy} E$ that is treated in an implicit way:

$$\frac{E^{n+1} - E^{n+1/2}}{\Delta z/2} = \frac{i}{2k_r \Delta x^2} \tilde{\Delta}_{xx} E^{n+1/2} + \frac{i}{2k_r \Delta y^2} \Delta_{yy} E^{n+1} + \frac{ik_0^2(\epsilon - n_r^2)}{4k_r} \left(E^{n+1} + E^{n+1/2} \right) . \tag{10.45}$$

It should be evident that this is only a minor modification to the simplest ADI method discussed previously. However, its coding may call for pre-calculated auxiliaries to hold all the position-dependent coefficients that appear in these equations. These will make it less error-prone to prepare the vectors that are passed to each tri-diagonal solve.

Collecting knowns and unknowns on the right and left, respectively, gives the two-stage update in a compact "vector form,"

$$(1 - i\delta_x \tilde{\Delta}_{xx} - i\gamma)E^{n+1/2} = (1 + i\delta_y \Delta_{yy} + i\gamma)E^n \quad , \tag{10.46}$$

$$(1 - i\delta_y \Delta_{yy} - i\gamma)E^{n+1} \quad = (1 + i\delta_x \tilde{\Delta}_{xx} + i\gamma)E^{n+1/2} \ , \tag{10.47}$$

where

$$\delta_x = \frac{\Delta z}{4k_r \Delta x^2} \quad , \quad \delta_y = \frac{\Delta z}{4k_r \Delta y^2} \quad , \quad \gamma_{i,j} = \frac{ik_0^2 \Delta z(\epsilon_{i,j} - n_r^2)}{4k_r} \ . \tag{10.48}$$

This can be written in an equivalent, but perhaps more explicit component form. The two stages read like this:

1. For each $y-$index $j = 1, \ldots, N_y$, solve this tri-diagonal linear system, for the vector of unknowns $X_k \equiv E_{k,j}^{n+1/2}$

$$\sum_k^{N_x} (\delta_{ik} - i\delta_x(\tilde{\Delta}_{xx})_{ik,j} - i\delta_{ik}\gamma_{i,j})X_k = (1 + i\gamma_{i,j})E_{i,j}^n + i\delta_y \sum_{n=j-1}^{j+1} (\Delta_{yy})_{jn} E_{i,n}^n \quad , \quad i = 1, \ldots, N_x$$

$$E_{k,j}^{n+1/2} = X_k \quad , \quad k = 1, \ldots, N_x \tag{10.49}$$

2. For each $x-$index $i = 1, \ldots, N_x$, solve this tri-diagonal linear system, for the vector of unknowns $Y_k \equiv E_{i,k}^{n+1}$

$$\sum_k^{N_y} (\delta_{jk} - i\delta_y(\Delta_{yy})_{jk} - i\delta_{jk}\gamma_{i,j})Y_k = (1 + i\gamma_{i,j})E_{i,j}^{n+1/2} + i\delta_x \sum_{n=i-1}^{i+1} (\tilde{\Delta}_{xx})_{in,j} E_{n,j}^{n+1/2} \quad , \quad j = 1, \ldots, N_y$$

$$E_{i,k}^{n+1} = Y_k \quad , \quad k = 1, \ldots, N_y \tag{10.50}$$

In comparison to the straightforward one-dimensional C-N update, the second step has the same structure of the vectors that are passed to the tri-diag solver as the three diagonals, but there is a slight difference in calculation of the right-hand-side vector. Its "non-diagonal" part "reaches" for neighbor field samples in the x-direction. This is also true for the right hand side in step one, neighbor values are fetched from the orthogonal direction. Perhaps the biggest change lies in preparing the linear system in the firs step, where different diagonals must be calculated for linear solves at different locations along dimension y.

10.5 Vectorial BPM

10.5.1 2D BPM for structured media using an iterated C-N Method

Iterated Crank-Nicolson Method is motivated by the difficulty in efficient application of the unmodified C-N approach to two dimensions. Indeed, with the tri-diagonal form of the matrix lost in two dimensions, the method looses lot of its performance. In the iterated C-N method, implicitness is given up in favor of the simple algorithm that tries to approximate the yet-to-be-calculated field samples in an iterative procedure.

The reader should notice that what is described in this section is in fact one special case of Method of Lines with a fixed ODE-solver method (and also with a fixed integration step — it is up to the user to ensure that accuracy is sufficient).

In the previous section we have seen that the evolution system for the transverse field components can be written in an abstract form

$$\partial_z \mathcal{E} = \hat{A}\mathcal{E}$$

where $\mathcal{E}$ stands for a vector of field samples composed of both E_x and E_y, and operator A is a "super-matrix" with elements $A_{xx}, A_{xy}, A_{yx}, A_{yy}$.

What one desires in the traditional implicit method is

$$\frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{\delta z} = \hat{A}\mathcal{E}^{n+1/2}$$

where $\mathcal{E}^{n+1/2}$ is approximated by averaging between current and previous z-step sample. Avoiding the need to solve big linear system, one forfeits implicitness and chooses to estimate $\mathcal{E}^{n+1/2}$ in a series of iterations as follows.

1. Execute explicit Euler step:
$$\frac{{}^{(1)}\mathcal{E}^{n+1} - \mathcal{E}^n}{\Delta z} = \hat{A}\mathcal{E}^n$$
2. Estimate the field at the midstep
$$\mathcal{E}^{n+1/2} = a_1\, {}^{(1)}\mathcal{E}^{n+1} + (1 - a_1)\mathcal{E}^n$$
3. Repeat the first step with the estimated RHS:
$$\frac{{}^{(2)}\mathcal{E}^{n+1} - \mathcal{E}^n}{\Delta z} = \hat{A}\, {}^{(1)}\mathcal{E}^{n+1/2}$$
4. Repeat averaging to improve the estimated midstep
$${}^{(2)}\mathcal{E}^{n+1/2} = a_2\, {}^{(2)}\mathcal{E}^{n+1} + (1 - a_2)\mathcal{E}^n$$
5. Finally, assume the above is a good estimate of the required midstep value, and execute Crank-Nicolson type step:
$$\frac{\mathcal{E}^{n+1} - \mathcal{E}^n}{\Delta z} = \hat{A}\, {}^{(2)}\mathcal{E}^{n+1/2}$$

All this method requires is an algorithm to calculate the action of operator $\hat{A}$ on an arbitrary vector of field-amplitude samples. This is relatively simple to implement with the help of expressions obtained in previous sections. The algorithm has also other attractive properties. It only requires "local" calculations which can be efficiently parallelized. Because the method is essentially explicit (or semi-implicit), it requires an integration step that is shorter than that in the full 2D Crank-Nicolson. However, since each step is relatively inexpensive to calculate, the overall speed of this approach is fully competitive with implicit methods.

10.5.2 Fully vectorial ADI method

Application of ADI method to a fully vectorial BPM problem requires little more care, and its implementation is more involved. The purpose of this section is to summarize the method in brief terms.

Let us order the transverse polarization components of the electric field into a two-part column vector

$$E = \begin{bmatrix} E_x \\ E_y \end{bmatrix} , \tag{10.51}$$

so that the propagation equations in the corresponding "matrix" form read

$$\partial_z E = i\hat{A}E . \tag{10.52}$$

Here, $\hat{A}$ is a two by two matrix operator with the familiar form

$$\hat{A} = \begin{bmatrix} A_{xx} & A_{xy} \\ A_{yx} & A_{yy} \end{bmatrix} \tag{10.53}$$

for which explicit expressions were derived earlier in this chapter. To apply the ADI scheme, one needs to decide on splitting the above operator into two portions, such that each can be treated implicitly along one coordinate direction. Moreover, we require that each operator component which will lead to a linear system only couples computational grid points in one coordinate direction. One such splitting can be written as follows:

$$\hat{A} = \hat{A}_1 + \hat{A}_2 = \begin{bmatrix} A^{(1)}_{xx} & 0 \\ A^{(1)}_{yx} & A^{(1)}_{yy} \end{bmatrix} + \begin{bmatrix} A^{(2)}_{xx} & A^{(2)}_{xy} \\ 0 & A^{(2)}_{yy} \end{bmatrix} \tag{10.54}$$

where

$$\begin{aligned}
\hat{A}^{(1)}_{xx}\mathcal{E}_x &= \frac{1}{2k_{ref}}\left[\frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2\mathcal{E}_x)\right] + \frac{1}{2}k_0^2(n^2 - n^2_{ref})\mathcal{E}_x\right] \\
\hat{A}^{(2)}_{xx}\mathcal{E}_x &= \frac{1}{2k_{ref}}\left[\frac{\partial^2\mathcal{E}_x}{\partial y^2} + \frac{1}{2}k_0^2(n^2 - n^2_{ref})\mathcal{E}_x\right] \\
\hat{A}^{(1)}_{yy}\mathcal{E}_y &= \frac{1}{2k_{ref}}\left[\frac{\partial^2\mathcal{E}_y}{\partial x^2} + \frac{1}{2}k_0^2(n^2 - n^2_{ref})\mathcal{E}_y\right] \\
\hat{A}^{(2)}_{yy}\mathcal{E}_y &= \frac{i}{2k_{ref}}\left[\frac{\partial}{\partial y}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2\mathcal{E}_y)\right] + \frac{1}{2}k_0^2(n^2 - n^2_{ref})\mathcal{E}_y\right] \\
\hat{A}^{(2)}_{xy}\mathcal{E}_y &= \frac{1}{2k_{ref}}\left[\frac{\partial}{\partial x}\left[\frac{1}{n^2}\frac{\partial}{\partial y}(n^2\mathcal{E}_y)\right] - \frac{\partial^2\mathcal{E}_y}{\partial x\partial y}\right] \\
\hat{A}^{(1)}_{yx}\mathcal{E}_x &= \frac{1}{2k_{ref}}\left[\frac{\partial}{\partial y}\left[\frac{1}{n^2}\frac{\partial}{\partial x}(n^2\mathcal{E}_x)\right] - \frac{\partial^2\mathcal{E}_x}{\partial y\partial x}\right]
\end{aligned} \tag{10.55}$$

Without going into details, we can generalize previously derived formula for the ADI method (9.19) and write the update as follows

$$E^{n+1} = \frac{1}{(1 - i\delta\hat{A}_1)}\frac{(1 + i\delta\hat{A}_2)}{(1 - i\delta\hat{A}_2)}\frac{(1 + i\delta\hat{A}_1)}{1}E^n \quad , \quad \delta = \Delta z/4 \tag{10.56}$$

Execution of this scheme requires that the inverse operator actions can be solved as a linear system with tri-diagonal matrices. For example, the last stage of the calculation will involve an "intermediate" vector X, on which an inverse operator must be applied:

$$Y = \frac{1}{(1 - i\delta\hat{A}_1)}X . \tag{10.57}$$

This is written as a linear system to be solved

$$(1 - i\delta\hat{A}_1)Y = X \ . \tag{10.58}$$

In the matrix form,

$$\begin{bmatrix} 1 - i\delta A_{xx}^{(1)} & -i\delta A_{xy}^{(1)} \\ 0 & 1 - i\delta A_{yy}^{(1)} \end{bmatrix} \begin{bmatrix} Y_x \\ Y_y \end{bmatrix} = \begin{bmatrix} X_x \\ X_y \end{bmatrix} \tag{10.59}$$

it becomes evident that this can be done in two steps. First, one takes the second row and solves for Y_y

$$(1 - i\delta A_{yy}^{(1)})Y_y = X_y \ . \tag{10.60}$$

Having found Y_y, the first row is used to solve for Y_x

$$(1 - i\delta A_{xx}^{(1)})Y_x = X_x + i\delta A_{xy}^{(1)} Y_y \ . \tag{10.61}$$

These two operations only require linear system solves that have tri-diagonal matrices generated by $A_{yy}^{(1)}$ and $A_{yy}^{(1)}$ in complete analogy to the semi-vectorial ADI-BPM. The only place where we encounter mixed derivatives is the last term, but this is a simple matrix-vector multiplication and as such it poses no problem. Naturally, the inverse $(1 + i\delta\hat{A}_2)^{-1}$ can be treated the same way. All put together, the fully vectorial ADI method too can be implemented such that the most "difficult" operation is the solution of a tri-diagonal linear system. As in the semi-vectorial case, these operations can be executed in parallel across one of the transverse dimensions of the computational domain.