Efficient solution of Maxwell's equations for geometries with repeating patterns by an exchange of discretization directions in the aperiodic Fourier modal method

Citation for published version (APA):

Document status and date:
Published: 01/01/2012

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
https://www.tue.nl/index.php?id=71870

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 13. Feb. 2019
Efficient solution of Maxwell’s equations for geometries with repeating patterns by an exchange of discretization directions in the aperiodic Fourier modal method

by

M. Pisarenco, J.M.L. Maubach, I.D. Setija, R.M.M. Mattheij
Efficient solution of Maxwell’s equations for geometries with repeating patterns by an exchange of discretization directions in the aperiodic Fourier modal method

M. Pisarenco\textsuperscript{a,b}, J.M.L. Maubach\textsuperscript{a}, I.D. Setija\textsuperscript{b}, R.M.M. Mattheij\textsuperscript{a}

\textsuperscript{a}Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600MB Eindhoven, The Netherlands
\textsuperscript{b}Department of Research, ASML Netherlands B.V., De Run 6501, 5504DR Veldhoven, The Netherlands

Abstract

The aperiodic Fourier modal method in contrast-field formulation is a numerical discretization and solution technique for solving scattering problems in electromagnetics. Typically, spectral discretization is used in the finite periodic direction and spatial discretization in the orthogonal direction. In the light of the fact that the structures of interest often have a large width-to-height ratio and that the two discretization approaches have different computational complexities, we propose exchanging the directions for spatial and spectral discretization. Moreover, if the scatterer has repeating patterns, swapping the discretization directions facilitates the reuse of previous computations. Therefore, the new method is suited for scattering from objects with a finite number of periods, such as gratings, memory arrays, metamaterials, etc. Numerical experiments demonstrate a considerable reduction of the computational costs in terms of time and memory. For a specific test case considered in this paper, the new method (based on alternative discretization) is 40 times faster and requires 100 times less memory than the method based on classical discretization.

Keywords: aperiodic Fourier modal method, AFMM-CFF, rigorous coupled-wave analysis, RCWA, electromagnetic scattering, perfectly matched layer, PML, computational complexity, alternative discretization
1. Introduction

Electromagnetic scattering occurs whenever propagating waves encounter obstacles on their path. It may be observed in various locations and at various lengthscales; from scattering of sunlight from the ocean surface to diffraction of a laser beam by a nanostructure. Maxwell’s equations define a mathematical model which rigorously describes the behavior of electromagnetic fields. The number of situations where an analytical solution of Maxwell equations can be found is very limited. Ref. (Bowman et al., 1987) gives a thorough overview of such very special cases. In all other cases numerical methods must be used to get an approximate solution of the equations. During the last decades many numerical methods for solving Maxwell equations have been developed. Several of them are widely used: the finite-element method (FEM) (Monk, 1992; Zschiedrich et al., 2006; Monk, 2003), the finite-difference time-domain method (FDTD) (Yee, 1966; Prather and Shi, 1999; Taflove and Hagness, 2005) and the integral equation methods (IEM), which include the boundary element method (BEM) (Prather et al., 1997; Li, 2010) and the volume integral method (VIM) (Zwamborn and van den Berg, 1991; Botha, 2006; Chang et al., 2006; Shcherbakov and Tishchenko, 2010; van Beurden, 2011).

The Fourier modal method (FMM) and the aperiodic Fourier modal method in contrast field formulation (AFMM-CFF) are numerical solution methods for Maxwell’s equations which are less known to the large audience in the field of computational electromagnetics. The FMM (Moharam et al., 1995a) is the oldest of the two and is also known under the name of rigorous coupled-wave analysis (RCWA). It is applicable for infinitely periodic structures and originated in the diffractive optics community more than 30 years ago (Knop, 1978). In the past decades the method has matured due to fundamental studies and improvements to its stability (Moharam et al., 1995b; Li, 1996a) and convergence (Granet and Guizal, 1996; Lalanne and Morris, 1996; Li, 1996b). Other important contributions to the evolution of the method are the techniques of adaptive spatial resolution (Granet, 1999) and normal vector fields (Popov and Nevière, 2000, 2001; Schuster et al., 2007). Ref. (Hench and Strakos, 2008) gives a mathematical perspective of the challenges that have been overcome in the FMM and of the open problems still to be addressed.

The AFMM-CFF (Pisarenco et al., 2010, 2011) is a recent extension of the FMM to aperiodic structures. It builds upon the aperiodic Fourier modal
method developed and refined by Lalanne and co-workers (Lalanne and Silberstein, 2000; Silberstein et al., 2001; Cao et al., 2002; Hugonin and Lalanne, 2005; Lecamp et al., 2007) with the key difference that the AFMM-CFF solves Maxwell’s equations formulated in terms of a contrast (scattered) field instead of a total field. This reformulation allows prescription of arbitrary incoming fields onto the structure of interest (Pisarenco et al., 2010). Aperiodicity is achieved by using perfectly matched layers (PMLs) (Berenger, 1994) on the vertical boundaries in order to annihilate the periodic boundary condition. An ideal PML yields an effective radiation boundary condition. Thus, unlike in the (periodic) FMM, effectively the same boundary condition is imposed on all boundaries. This fact facilitates the exchange of discretization directions and, as shown later, decreases the computational costs.

Figure 1 shows a two-dimensional cylinder problem discretized by several numerical methods. From the point of view of computational costs, the discretization process determines the number of degrees of freedom of the discrete problem. This number in turn determines the execution time and the memory requirements. For instance, a discretization of Maxwell equations (containing six unknowns) on a regular grid as in FDTD yields $N_{\text{DOF}} = 6NM$ degrees of freedom, where $N$ and $M$ represent the number of discretization points in the $x$- and $z$-direction respectively. If the system is subsequently solved by an iterative method the computational cost is typically $O(N_{\text{DOF}}^2)$ =
\(O(N^2M^2)\). Similar reasoning holds for the FEM and the integral equation methods, although the grid may be irregular in these cases. The key point here is that the computational cost of the methods mentioned above scales in the same way for all the discretization directions.

As can be seen from Figure 1, the FMM and AFMM-CFF make a distinction between the directions. In the vertical direction (the \(z\)-direction) the domain is divided into \(M\) layers in which the material properties are assumed to be \(z\)-independent. In each layer, the horizontal direction (the \(x\)-direction) is discretized by applying the Galerkin method with ”shifted” Fourier harmonics as basis and test functions. This yields a system of ordinary differential equations of a size equal to the number of basis functions, \(N\). The general solution (with unknown integration constants) of this system may be derived by numerically solving a matrix diagonalization problem. The diagonalization of dense matrices (based on QR-decomposition) has a computational complexity of \(O(N^3)\) (Golub and Van Loan, 1996, Section 5.2). Thus, for \(M\) layers the total computational cost is \(O(N^3M)\). Similarly the memory requirements can be estimated to \(O(N^2M)\).

It appears that the \(x\)-direction is ”more expensive” both in terms of time and memory. For rectangular scatterers/domains that are much longer in the \(x\)-direction, it is reasonable to choose an alternative discretization: make the longer direction ”cheaper” by using spatial discretization into layers and apply spectral discretization in the shorter direction. This exchange of directions turns out to be possible for the AFMM-CFF and not for the FMM since the latter requires different boundary conditions on the vertical and horizontal boundaries. For structures with repeating patterns, the alternative discretization introduces slicing in the direction of repetition. We exploit this fact by proposing a new recursive algorithm which steps through the slices in a geometric progression. This further reduces the computational costs by replacing the linear dependence of the costs on the number of slices by a logarithmic one.

In many important applications the scatterer is placed in a stratified medium. Examples include broad scientific areas such as submarine detection, geophysical exploration, optical microscopy and metrology (Li, 2010). Therefore, in this paper we consider the more general case of a bounded scatterer placed in a stratified medium, also referred to as background multilayer.

This paper is structured as follows. In the next section we formulate a horizontal and an equivalent vertical problem whose solutions are to be computed with the AFMM-CFF using respectively classical and alternative
discretizations. Section 3 contains a general formulation of the AFMM-CFF suitable for both problems defined previously. The attention here is on the computation of the background field for the vertical problem. In Section 4 we specifically consider structures with repeating patterns and present a fast recursive algorithm for solving the final linear system exploiting this property. A summary of the theoretical estimates of time and memory required by the AFMM-CFF is given in Section 5. Section 6 presents numerical evidence on the efficiency of the proposed approach. Experimental speed-up factors are compared to the theoretical ones. Finally, our conclusions are presented in Section 7.

2. Problem formulation

We consider electromagnetic fields with a time-harmonic behavior, i.e.

\[ \tilde{e}(x, t) = Re\{e(x)e^{i\omega t}\}, \]
\[ \tilde{h}(x, t) = Re\{h(x)e^{i\omega t}\}, \]

where \( \omega \) is the angular temporal frequency. Using the time-harmonic property of the fields, Maxwell’s equations transform to

\[ \nabla \times e(x) = -k_0 h(x), \]
\[ \nabla \times h(x) = -k_0 \epsilon(x, z) e(x). \]

Here \( k_0 = \omega \sqrt{\epsilon_0 \mu_0} \), \( x = (x, y, z)^T \) is the position vector, \( e = (e_x, e_y, e_z)^T \) is the electric field and \( h = (h_x, h_y, h_z)^T \) is the magnetic field scaled by \(-i \sqrt{\epsilon_0 / \mu_0}\) (with respect to the physical magnetic field). We assume that all materials are non-magnetic and that the electric permittivity \( \epsilon \) is a scalar \( y \)-invariant function. The incident field is given by

\[ e^{inc}(x) = a e^{-ik^{inc} \cdot x}, \]

where \( k^{inc} = (k_{inc}^x, k_{inc}^y, k_{inc}^z)^T \) is the wavevector and \( a = (a_x, a_y, a_z)^T \) is the amplitude vector. Note that Maxwell’s equations in homogeneous isotropic media require that \( k^{inc} \cdot a = 0 \).

The numerical solution of (3) is to be computed on a domain \( \Omega = [0, \Lambda] \times \mathbb{R} \times \mathbb{R} \) using AFMM-CFF with classical and alternative discretization. Note
that alternative discretization may be applied by a "turn" of the coordinate system

\[(x, y, z)^T \rightarrow (z, -y, x)^T, \quad (4)\]
such that formally the discretization directions remain unchanged, i.e. the \(x\)-direction is discretized with harmonics and the \(z\)-direction is discretized with slices in both approaches. Figure 2 demonstrates the turn for a finite grating with three rectangular lines. Note that instead of three slices, we now have seven. From now on, we distinguish a horizontal problem \(\bar{P}\) consisting of (3) with a geometry \(\epsilon(x, z) = \bar{\epsilon}(x, z)\) and an incoming field \(\bar{e}^{\text{inc}}(x) = \bar{a} e^{-i\bar{k}^{\text{inc}} \cdot x}\) and a vertical problem \(\bar{\bar{P}}\) with \(\epsilon(x, z) = \bar{\bar{\epsilon}}(x, z)\) and \(\bar{\bar{e}}^{\text{inc}}(x) = \bar{\bar{a}} e^{-i\bar{\bar{k}}^{\text{inc}} \cdot x}\). We define a transformation matrix

\[
T = \begin{pmatrix}
0 & 0 & 1 \\
0 & -1 & 0 \\
1 & 0 & 0
\end{pmatrix},
\]

which realizes the rotation of the problem. Note that a second application of the transformation yields the original coordinate system, i.e. \(TT = I\) or \(T = T^{-1}\). The vertical problem is related to the horizontal problem through

\[
\bar{\bar{\epsilon}}(x) = \bar{\epsilon}(Tx),
\]

\[
\bar{\bar{a}} = Ta,
\]

\[
\bar{\bar{k}}^{\text{inc}} = Tk^{\text{inc}},
\]
and the solutions of the two problems satisfy

\[
\bar{e}(x) = T\bar{e}(Tx), \quad (7a)
\]
\[
\bar{h}(x) = T\bar{h}(Tx). \quad (7b)
\]

The functions \( \bar{e} \) and \( \bar{h} \) determine the geometry of the problem which typically consists of a scatterer placed on a substrate (see Figure 2). In many applications the scatterer has a width which is much larger than its height which motivates the turn of coordinates. Although the incident field is assumed to be a plane wave in (3c), any other wavefront may be modeled by a superposition of multiple plane waves.

In the next section we describe a generalized AFMM-CFF which can be applied to both horizontal and vertical problems. Where necessary we discuss the differences and modifications required in order to solve the vertical problem \( \bar{P} \).

3. Generalized AFMM-CFF for horizontal and vertical backgrounds

3.1. Perfectly matched layers and the Maxwell equations for the contrast field

We use PMLs (Berenger, 1994) in order to impose the radiation condition at the lateral boundaries. The radiation condition however imposes a restriction on the problems which can be solved: no incident field is allowed. Therefore, Maxwell’s equations need to be reformulated such that the incident field is replaced by an equivalent source term. To this end, an associated background problem is defined,

\[
\nabla \times e^b(x) = -k_0h^b(x), \quad (8a)
\]
\[
\nabla \times h^b(x) = -k_0\epsilon^b(x, z)e^b(x), \quad (8b)
\]

with

\[
e^{b,inc}(x) = ae^{-ik^{inc}x}. \quad (8c)
\]

This problem is chosen such that it admits an analytical solution and the function \( \epsilon - \epsilon^b \) has compact support. As explained in (Pisarenco et al., 2010), the compact support condition is required in order to avoid non-zero source terms in the PML. Typically \( \epsilon^b \) represents the permittivity of the background multilayer which supports the scatterer. It may be horizontal or
vertical, depending on the problem being solved. Subtraction of (8) from (3) yields the so-called contrast-field formulation

$$\nabla \times \mathbf{e}^c(x) = -k_0 h^c(x),$$

$$\nabla \times h^c(x) = -k_0 \epsilon(x, z) e^c(x) - k_0 (\epsilon(x, z) - \epsilon^b(x, z)) e^b(x),$$

with

$$e^{c,inc}(x) = 0.$$

The PMLs can be viewed as an analytical continuation of the solution into the complex plane (Chew and Weeden, 1994; Collino and Monk, 1998b). For PMLs placed in the $x$-direction, this implies a change of the $x$-derivative in the differential equations (9):

$$\frac{\partial}{\partial x} \rightarrow \frac{1}{f'(x)} \frac{\partial}{\partial x}, \text{ with } f(x) = x + i\beta(x).$$

The function $\beta$ is continuous and non-zero only in the PMLs, which are placed in the stripes $[0, x_l]$ and $[x_r, \Lambda]$. We mention that although PMLs are originally designed for hyperbolic equations, they have also been recently successfully applied to the heat equation (Lantos and Nataf, 2010). Lalanne and co-workers (Lalanne and Silberstein, 2000; Silberstein et al., 2001; Hugo-nin and Lalanne, 2005) were the first to use absorbing layers and PMLs with
Fourier modal methods. In (Hugonin and Lalanne, 2005) they chose trigono-
metric stretching functions whose Fourier coefficients can be computed ana-
lytically. Many other forms for $f$ (and implicitly for $\beta$) have been suggested
(Berenger, 1996; Collino and Monk, 1998a; Petropoulos, 2003). The most
successful use a polynomial or geometric variation of $f$ in the PML. We
adopt the first form. An example of such a function is shown in Figure 3.

We assume a PML given by the coordinate transformation

$$f(x) = x + i\beta(x) = x + ik_0p(\beta_0|x - x_0|)^{p+1}, \quad p \in \mathbb{R}^+,$$

(11)

where $x_0 = x_l$ for the left PML and $x_0 = x_r$ for the right PML. To the
authors’ knowledge, no study exists on the choice of $p$ for PMLs in the
aperiodic Fourier modal methods. For instance, for FDTD methods nearly
optimal results have been obtained for $p \in [3, 4]$ (Berenger, 1996; Wu and
Fang, 1996).

We look at the amplitude of a plane wave with unit amplitude after
passing through a PML of width $\Delta x$ (see Figure 4)

$$|e^{-ik_x f(x)}| = |e^{-ik_x x}| |e^{-k_x k_0^p(\beta_0\Delta x)^{p+1}}| \leq e^{-(k_0\beta_0\Delta x)^{p+1}}.$$  

Now, we determine the distance over which the amplitude decays from 1 to $e^{-1}$.

$$-(k_0\beta_0\Delta x)^{p+1} = -1 \Rightarrow k_0\Delta x = \frac{1}{\beta_0}.$$
This relation provides an intuitive meaning for $\beta_0$ - it determines the inverse decay length. Since $k_0 = 2\pi/\Lambda$, the length is scaled by the wavelength of the incident field.

3.2. Discretization

Spatial discretization in $z$ is achieved in the FMM by dividing the computational domain vertically into $M$ slices such that the permittivity may be considered $z$-independent in each separate slice. The terminology of ”slice” is preferred here to ”layer”, as the latter is reserved to physical layers in the multilayer stack. As illustrated in Figure 5, the upper and lower interface of slice $l$ are located at $h_{l-1}$ and $h_l$ respectively. For future reference, we define the set $h = \{h_l, l = 0, ..., M\}$. Since $z \in \mathbb{R}$, we take $h_0 = -\infty$, $h_M = +\infty$. The permittivities in each slice are given by

$$\epsilon_l(x) = \epsilon(x, z_l), \text{ with } z_l \in [h_{l-1}, h_l),$$  \hspace{1cm} (12)

$$\epsilon^b_l(x) = \epsilon^b(x, z_l), \text{ with } z_l \in [h_{l-1}, h_l).$$  \hspace{1cm} (13)

Thus, the profile of the scatterer is approximated by a staircase as in Figure 5. In the horizontal problem, the permittivity of the background $\epsilon^b_0$ is constant in each layer. The electric and magnetic fields on the computational domain now consist of fields in separate slices

$$\mathbf{e}_l(x, y, z) = \mathbf{e}(x, y, z), \text{ } z \in [h_{l-1}, h_l),$$  \hspace{1cm} (14)

$$\mathbf{h}_l(x, y, z) = \mathbf{h}(x, y, z), \text{ } z \in [h_{l-1}, h_l).$$  \hspace{1cm} (15)
The Maxwell equations for the contrast field (9) with PMLs (11) applied to slice $l$ read

\[
\frac{\partial}{\partial y} e_{z,l} - \frac{\partial}{\partial z} e_{y,l} = -k_0 h_{x,l} 
\]

\[
\frac{\partial}{\partial z} e_{x,l} - \frac{\partial}{\partial x} e_{y,l} = -k_0 h_{z,l} 
\]

\[
\frac{1}{f'(x)} \frac{\partial}{\partial x} e_{x,l} - \frac{\partial}{\partial y} e_{z,l} = -k_0 h_{y,l} 
\]

\[
\frac{1}{f'(x)} \frac{\partial}{\partial y} h_{z,l} - \frac{\partial}{\partial x} h_{y,l} = -k_0 e_{c,x,l} - k_0 (1 - \frac{1}{\epsilon_l(x)}) e_{b,x,l} 
\]

\[
\frac{\partial}{\partial x} h_{y,l} - \frac{\partial}{\partial z} e_{x,l} = -k_0 \epsilon_l(x) e_{c,z,l} - k_0 (\epsilon_l(x) - \epsilon_b(x)) e_{b,z,l} 
\]

for $(x, y, z) \in [0, \Lambda] \times \mathbb{R} \times [h_{l-1}, h_l)$. Equation (16d) has been divided by $\epsilon_l(x)$ in order to avoid products of functions with concurrent (in the same point) jump discontinuities on the right-hand side. As shown by Li (1996b), discretization of the original equation leads to slower convergence.

Spectral discretization in $x$ is achieved with a Galerkin approach that uses "shifted" Fourier harmonics as basis functions and test functions,

\[
\phi_n(x, y) = e^{-i(k_{xn}x + k_{yn}y)}, \quad (17)
\]

where \(k_{xn} = k_{x}^{inc} + n \frac{2\pi}{\Lambda}, \quad k_{yn} = k_{y}^{inc}, \quad \text{for } n = -N... + N.\)

In each slice $l$ the contrast (electric and magnetic) fields are expanded as

\[
e_{\alpha,l}^{c,N}(x, y, z) = \sum_{n=-N}^{N} s_{\alpha,l,n}(z) \phi_n(x, y) = (s_{\alpha,l}(z))^T \cdot \phi(x, y), \quad (18a)
\]

\[
h_{\alpha,l}^{c,N}(x, y, z) = \sum_{n=-N}^{N} u_{\alpha,l,n}(z) \phi_n(x, y) = (u_{\alpha,l}(z))^T \cdot \phi(x, y). \quad (18b)
\]

The $\alpha$ symbol stands for the $x$-, $y$-, or $z$-component of the field. The background fields, which are known in advance, have to be represented in the
same basis as the contrast field, i.e.
\[ e_{a,l}^{b,N}(x, y, z) = \sum_{n=-N}^{N} s_{a,l,n}^{b}(z) \phi_n(x, y) = (s_{a,l}^{b}(z))^T \cdot \phi(x, y), \quad (19a) \]
\[ h_{a,l}^{b,N}(x, y, z) = \sum_{n=-N}^{N} u_{a,l,n}^{b}(z) \phi_n(x, y) = (u_{a,l}^{b}(z))^T \cdot \phi(x, y). \quad (19b) \]

Computation of \( s_{a,l,n}^{b}(z) \) and \( u_{a,l,n}^{b}(z) \) for both horizontal and vertical multi-layers will be discussed in Section 3.3.

We apply the Galerkin method with a standard inner product on the interval \([0, \Lambda)\) to the contrast field equations (16) to obtain a system of ordinary differential equations in \( z \) for each slice \( l \),

\[-iK_x s_{x,l}^{c}(z) - k_0^{-1} \frac{d}{dz} s_{y,l}^{c}(z) = -u_{y,l}^{c}(z), \quad (20a)\]
\[ k_0^{-1} \frac{d}{dz} s_{x,l}^{c}(z) + iFK_x s_{x,l}^{c}(z) = -u_{y,l}^{c}(z), \quad (20b)\]
\[-iFK_x s_{y,l}^{c}(z) + iK_y s_{x,l}^{c}(z) = -u_{y,l}^{c}(z), \quad (20c)\]
\[-iK_y u_{x,l}^{c}(z) - k_0^{-1} \frac{d}{dz} u_{y,l}^{c}(z) = -P_{l}^{-1} s_{x,l}^{c}(z) - (P_{l}^{-1} - P_{l}^{b-1}) s_{y,l}^{b}(z), \quad (20d)\]
\[ k_0^{-1} \frac{d}{dz} u_{x,l}^{c}(z) + iFK_x u_{x,l}^{c}(z) = -E_{l} s_{y,l}^{c}(z) - (E_{l} - E_{l}^{b}) s_{y,l}^{b}(z), \quad (20e)\]
\[-iFK_x u_{y,l}^{c}(z) + iK_y u_{x,l}^{c}(z) = -E_{l} s_{x,l}^{c}(z) - (E_{l} - E_{l}^{b}) s_{x,l}^{b}(z). \quad (20f)\]

Introducing the notation for Fourier coefficients \( \hat{\xi}_n \) of a function \( \xi(x) \) on the interval \( x \in [0, \Lambda) \),

\[ \hat{\xi}_n = \int_{0}^{\Lambda} \xi(x) e^{-in\frac{2\pi}{\Lambda}x} \, dx, \]

the matrices in the expressions above are defined as follows,

\[ (K_x)_{mn} = (k_{xn}/k_0) \delta_{mn}, \quad (21a)\]
\[ (K_y)_{mn} = (k_{yn}/k_0) \delta_{mn}, \quad (21b)\]
\[ (E_l)_{mn} = \hat{\epsilon}_{l,n-m}, \quad (21c)\]
\[ (P_l)_{mn} = \hat{p}_{l,n-m}, \quad (21d)\]
\[ (E_l^{b})_{mn} = \hat{\epsilon}_{l,n-m}^{b}, \quad (21e)\]
\[ (P_l^{b})_{mn} = \hat{p}_{l,n-m}^{b}, \quad (21f)\]
\[ (F)_{mn} = \hat{\gamma}_{n-m}, \quad (21g)\]
for \( m, n = -N... + N \). Here \( \delta_{mn} \) is the Kronecker delta and

\[
\begin{align*}
p_t(x) &= 1/\epsilon_t(x) \quad (22a) \\
p_t^b(x) &= 1/\epsilon_t^b(x) \quad (22b) \\
\gamma(x) &= 1/f'(x) \quad (22c)
\end{align*}
\]

The function \( f(x) \) is the complex coordinate transformation implementing the PML and has been defined in Section 3.1. Note that for a horizontal background the matrices \( E_t^b = \epsilon_t^b \mathbf{I} \) and \( P_t^b = (\epsilon_t^b)^{-1} \mathbf{I} \) are diagonal.

3.3. Discretization of the background field

The background field is the solution to the problem of light propagation in a multilayer stack. In the horizontal problem, we assume a multilayer stack which extends infinitely in the \( xy \)-plane and is invariant in the \( x \)- and \( y \)-directions. Let the locations of interfaces in the multilayer be given by the set \( h^b = \{ h_l^b, l = 0, ..., M^b \} \). Due to the nature of spatial discretization, we have \( h^b \subseteq h \). We will therefore decompose the background field depending on discretization slices instead of physical layers. In each slice the field consists of an upward- and downward-traveling wave.

\[
\begin{align*}
\mathbf{e}^b(x, y, z) &= \mathbf{t}_s, l e^{-i k_0^b (z - h_l^b)} e^{-i (k^\text{inc}_x x + k^\text{inc}_y y)} + \mathbf{r}_s, l e^{i k_0^b (z - h_l^b)} e^{-i (k^\text{inc}_x x + k^\text{inc}_y y)}, \quad (23a) \\
\mathbf{h}^b(x, y, z) &= \mathbf{t}_u, l e^{-i k_0^b (z - h_l^b)} e^{-i (k^\text{inc}_x x + k^\text{inc}_y y)} + \mathbf{r}_u, l e^{i k_0^b (z - h_l^b)} e^{-i (k^\text{inc}_x x + k^\text{inc}_y y)}, \quad (23b)
\end{align*}
\]

for \( z \in [h_{l-1}, h_l] \), where

\[
\bar{q}_l = i \sqrt{\frac{\epsilon_t^b}{\epsilon_t^b} - \left( \frac{k^\text{inc}_x}{k_0} \right)^2 - \left( \frac{k^\text{inc}_y}{k_0} \right)^2}, \quad (24)
\]

and the amplitude vectors \( \mathbf{t} \) and \( \mathbf{r} \) can be determined from the interface conditions resulting from the Maxwell equations, see (Yeh, 2005; Schädle et al., 2007).

In the vertical problem the multilayer is invariant in the \( z \)- and \( y \)-directions. Application of the transformation (7) to (23) yields the vertical background field.
Figure 6: The background field consisting of transmitted and reflected waves for the horizontal (left) and vertical (right) problem.

\[ e_b(x, y, z) = e^{-ik_{inc}z}e^{-k_0q_1(x-h_1)}e^{-ik_{inc}y} \]
\[ + \bar{r}_{s;I} e^{-ik_{inc}z}e^{k_0q_1(x-h_1)}e^{-ik_{inc}y}, \quad x \in [h_{l-1}, h_l), \]  
\[ h_b(x, y, z) = h^{-i\hat{k}_{inc}z}e^{k_0q_1(x-h_1)}e^{-i\hat{k}_{inc}y} \]
\[ + \bar{r}_{u;I} e^{-i\hat{k}_{inc}z}e^{k_0q_1(x-h_1)}e^{-i\hat{k}_{inc}y}, \quad x \in [h_{l-1}, h_l). \]  

where \( \bar{r}_{u;I} = T \bar{r}_{u;I} \), \( \bar{r}_{s;I} = T \bar{r}_{s;I} \), \( \bar{t}_{s;I} = T \bar{t}_{s;I} \), \( \bar{t}_{u;I} = T \bar{t}_{u;I} \). Figure 6 depicts the background fields for the horizontal and vertical problems. We need to write the background fields in a discretized form (19). It will be shown that the following less general form can be used for both settings (horizontal and vertical),

\[ e^b_\alpha(x, y, z) = (s^b_\alpha(z))^T \cdot \phi(x, y) = (d_{s,\alpha} s^b_\alpha(z))^T \cdot \phi(x, y), \]  
\[ h^b_\alpha(x, y, z) = (u^b_\alpha(z))^T \cdot \phi(x, y) = (d_{u,\alpha} u^b_\alpha(z))^T \cdot \phi(x, y). \]  

The horizontal background field can be exactly represented in this form. Since the \((x, y)\)-dependent part of the horizontal background field (23) coincides with \( \phi_0(x, y) \), we have \( d_{s,\alpha} = d_{u,\alpha} = e_{N+1} \), where \( e_{N+1} \in \mathbb{R}^{2N+1} \) is defined as \( (e_{N+1})_n = \delta_{n,N+1} \), and

\[ \bar{s}^b_\alpha(z) = \bar{s}_{s,\alpha;I} e^{-k_0q_1(z-h_1)} + \bar{s}_{s,\alpha;L} e^{k_0q_1(z-h_1)}, \quad z \in [h_{l-1}, h_l), \]  
\[ \bar{u}^b_\alpha(z) = \bar{u}_{u,\alpha;I} e^{-k_0q_1(z-h_1)} + \bar{u}_{u,\alpha;L} e^{k_0q_1(z-h_1)}, \quad z \in [h_{l-1}, h_l). \]  

On the other hand, the vertical background field can only be represented in the form (19) or (26) through a projection. The basis function \( \phi_0(x, 0) \)
gives the $x$-dependence for all plane waves constituting the background field in Figure 6 (left). None of the basis functions in the vertical problem give an exact representation of the background field in Figure 6 (right). In other words, in the vertical problem the basis used for the contrast field is not an eigenbasis for the background problem. We separate the $x$-dependent part of the vertical background field (25),

$$b_{s,a}(x) = \tilde{t}_{s,a} e^{-ik \bar{q}(x-h_i)} + \bar{r}_{s,a} e^{ik \bar{q}(x-h_i)}, \quad x \in [h_{i-1}, h_i), \quad (28a)$$

$$b_{u,a}(x) = \tilde{t}_{u,a} e^{-ik \bar{q}(x-h_i)} + \bar{r}_{u,a} e^{ik \bar{q}(x-h_i)}, \quad x \in [h_{i-1}, h_i), \quad (28b)$$

and project $b_{s/u,a}(x)$ on the available basis $\phi_n(x, 0) = e^{-ik_{zn} x}$,

$$b_{s/u,a}(x) = \sum_{n=-N}^{N} \tilde{b}_{s/u,a,n} e^{-ik_{zn} x}. \quad (29)$$

Multiplying both sides by $e^{ik_{zn} x}$ and integrating over the interval $[0, \Lambda]$, yields

$$\tilde{b}_{s/u,a,n} = \frac{1}{\Lambda} \int_{0}^{\Lambda} b_{s/u,a}(x)e^{ik_{zn} x} dx = \frac{1}{\Lambda} \int_{0}^{\Lambda} (b_{s/u,a}(x)e^{ik_{0z} x})e^{-in\frac{2\pi}{\Lambda} x} dx.$$

Thus the coefficients $\tilde{b}_{s/u,a,n}$ may be computed with the help of a Fourier transform (FFT). We denote by $b_{s/u,a}$ the vector formed from these coefficients. Now, the background field in the scatterer can be written in the available basis

$$\bar{e}_{\alpha}^{b,N} (x, y, z) = e^{-i\bar{k}_{zn} z} \sum_{n=-N}^{N} \tilde{b}_{s,a,n} \phi_n(x, y) = (b_{s,a} e^{-i\bar{k}_{zn} z})^T \phi(x, y), \quad (30)$$

$$\bar{h}_{\alpha}^{b,N} (x, y, z) = e^{-i\bar{k}_{zn} z} \sum_{n=-N}^{N} \tilde{b}_{u,a,n} \phi_n(x, y) = (b_{u,a} e^{-i\bar{k}_{zn} z})^T \phi(x, y). \quad (31)$$

The fields $\bar{e}_{\alpha}^{b,N}$, $\bar{h}_{\alpha}^{b,N}$ approximate the exact fields $\bar{e}_{\alpha}$, $\bar{h}_{\alpha}$ in (25) with $2N + 1$ basis functions. Thus in representation (26) we have $\bar{d}_{s,a} = b_{s,a}$, $\bar{d}_{u,a} = b_{u,a}$ and

$$\bar{s}_{\alpha}^b (z) = \bar{u}_{\alpha}^b (z) = e^{-i\bar{k}_{zn} z}. \quad (32)$$

To achieve uniform notation in the horizontal and vertical settings, we denote

$$\bar{q}_i = i\bar{k}_{zn}.$$

15
From (27) and (32) we conclude that the following differential equation is satisfied for \( z \in [h_{l-1}, h_l) \),

\[
\begin{align*}
\frac{d}{dz^2}s^b_\alpha & = q_l^2 s^b_\alpha, \\
\frac{d}{dz^2}u^b_\alpha & = q_l^2 u^b_\alpha.
\end{align*}
\]

(34a) (34b)

This property will be used in the next subsection.

3.4. General solution in a slice

The \( y \)- and \( z \)-component of the electric and magnetic field can be eliminated from the discretized Maxwell’s equations for the contrast field (20),

\[
\begin{align*}
\frac{d}{dz^2}s^c_{x,l}(z) & = k_0^2 C_l s^c_{x,l}(z) + k_0^2 (B_l P_l^{-1} - B^b_l (P^b_l)^{-1}) d_{s,\alpha} s^b_\alpha(z), \\
\frac{d}{dz^2}u^c_{x,l}(z) & = k_0^2 D_l u^c_{x,l}(z) - k_0^2 (E_l - E^b_l) d_{u,\alpha} u^b_\alpha(z),
\end{align*}
\]

(35a) (35b)

with

\[
\begin{align*}
A_l & = (F K_x)^2 - E_l, \\
B_l & = F K_x F^{-1} K_x - I, \\
B^b_l & = F K_x (E^b_l)^{-1} F K_x - I, \\
C_l & = K_g^2 + B_l P_l^{-1}, \\
D_l & = K^2 + A_l.
\end{align*}
\]

(36a) (36b) (36c) (36d) (36e)

In order to arrive at (35) from (20), additionally the Maxwell equations for the background field (8) have been used. The solution vector consists of a homogeneous and particular solution

\[
\begin{align*}
s^c_{x,l}(z) & = s^c_{x,\text{hom},l} + s^c_{x,\text{part},l}, \\
u^c_{x,l}(z) & = u^c_{x,\text{hom},l} + u^c_{x,\text{part},l}.
\end{align*}
\]

(37a) (37b)

The homogeneous solution is given by

\[
\begin{align*}
s^c_{x,\text{hom},l} & = W_{s,l} (e^{-k_0 Q_{s,l}(z-h_l)} c^+_{s,l} + e^{k_0 Q_{s,l}(z-h_{l+1})} c^-_{s,l}), \\
u^c_{x,\text{hom},l} & = W_{u,l} (e^{-k_0 Q_{u,l}(z-h_l)} c^+_{u,l} + e^{k_0 Q_{u,l}(z-h_{l+1})} c^-_{u,l}),
\end{align*}
\]

(38a) (38b)
where the pairs $W_{s,l}$, $Q_{s,l}$ and $W_{u,l}$, $Q_{u,l}$ contain the matrix of eigenvectors and the diagonal matrix of square roots of eigenvalues of, respectively, $C_l$ and $D_l$. The radiation condition is applied in the $z$-direction by allowing only outgoing waves in the top and bottom layers,

$$c_{s,1}^+ = c_{u,1}^+ = c_{s,M}^- = c_{u,M}^- = 0.$$  \hfill (39)

Both the PML and the above constraint implement the same effective boundary condition, such that the same problem is solved with classical and alternative discretizations. In the classical FMM, where a periodic boundary condition is imposed at two of the sides, the two discretizations solve different problems. This implies that the exchange of directions is only possible for finite structures, and not for infinitely periodic ones.

To find the particular solution we assume the form

$$s_{c,x;part,l}(z) = p_{s,l}s_{x}(z), \hfill (40a)$$

$$u_{c,x;part,l}(z) = p_{u,l}u_{x}(z). \hfill (40b)$$

The solution has this form for right-hand sides with a finite family of derivatives, such as polynomial and trigonometric functions (method of undetermined coefficients applied to systems). Substituting this ansatz in Equation (35) and using property (34), we obtain two linear systems which can be solved for $p_{s,l}$ and $p_{u,l}$

$$(C_l - q_l^2 I)p_{s,l} = -(B_lP_l^{-1} - B_l^p(P_l^p)^{-1})d_{s,x}, \hfill (41a)$$

$$(D_l - q_l^2 I)p_{u,l} = (E_l - E_l^p)d_{u,x}. \hfill (41b)$$

3.5. Matching of solutions at interfaces

At the interface, continuity of the tangential components of the contrast fields is required

$$\left( \epsilon_{l+1}^c - \epsilon_l^c \right) \times n = 0, \hfill (42a)$$

$$\left( h_{l+1}^c - h_l^c \right) \times n = 0. \hfill (42b)$$

These interface conditions follow from the continuity of the tangential components of the total and background fields. After applying the Galerkin
approach as described in Section 3.2 these equations become

\( s_{x,l}(h_{t+1}) = s_{x,l+1}(h_{t+1}), \quad (43a) \)

\( s_{y,l}(h_{t+1}) = s_{y,l+1}(h_{t+1}), \quad (43b) \)

\( u_{x,l}(h_{t+1}) = u_{x,l+1}(h_{t+1}), \quad (43c) \)

\( u_{y,l}(h_{t+1}) = u_{y,l+1}(h_{t+1}). \quad (43d) \)

Using (20c) in (20e) and (20f) in (20b), the \( y \)-components of the fields are expressed in terms of \( x \)-components,

\( s_{c,y,l}(h_{l+1}) = s_{c,y,l+1}(h_{l+1}), \quad (43a) \)

\( u_{c,y,l}(h_{l+1}) = u_{c,y,l+1}(h_{l+1}). \quad (43b) \)

We define

\[
W_l = \begin{bmatrix}
0 \\
A_l^{-1} W_{u,l} Q_{a,l} & A_l^{-1} F K_x K_y W_{s,l}
\end{bmatrix},
\]

\[
V_l = - \begin{bmatrix}
W_{u,l} \\
B_l^{-1} F K_x E_l^{-1} K_y W_{u,l} & B_l^{-1} W_{s,l} Q_{s,l}
\end{bmatrix},
\]

\[
X_l = \begin{bmatrix}
e^{-k_0 Q_{a,l}(h_{l+1} - h_l)} \\
e^{-k_0 Q_{a,l}(h_{l+1} - h_l)} 0
\end{bmatrix},
\]

and

\[
c_l^+ = \begin{bmatrix}
-c_{u,l}^+ \\
-c_{s,l}^+
\end{bmatrix}, \quad c_l^- = \begin{bmatrix}
c_{u,l}^- \\
c_{s,l}^-
\end{bmatrix}.
\]

Then, from (43), (37), (38), (40) we have for each slice

\[
\begin{bmatrix}
W_l X_l & W_l \\
V_l X_l & -V_l
\end{bmatrix} \begin{bmatrix}
c_l^+ \\
c_l^-
\end{bmatrix} + g_l(h_{l+1}) = \begin{bmatrix}
W_{l+1} X_{l+1} \\
V_{l+1} - V_{l+1} X_{l+1}
\end{bmatrix} \begin{bmatrix}
c_{l+1}^+ \\
c_{l+1}^-
\end{bmatrix} + g_{l+1}(h_{l+1}),
\]

where

\[
g_l(z) = \begin{bmatrix}
A_l^{-1}(F K_x K_y p_{s,l} s_b^b + k_0^{-1} p_{u,l} d_s^b) + (E_l - E_b) d_{s,y}^b \\
B_l^{-1}(F K_x E_l^{-1} K_y p_{u,l} u_b^b + k_0^{-1} p_{s,l} d_s^b - i F K_x (I - E_l^{-1} E_b) d_{s,z}^b)
\end{bmatrix}.
\]
In (47a) the vectors $c_l^+$, $c_l^-$ are unknown, except the vectors corresponding to incoming fields in the upper and bottom slices,

$$c_1^+ = 0, \quad c_M^- = 0.$$  \hfill (47b)

The coupled linear systems (47) for $l = 1, \ldots, M - 1$ have to be solved in a stable manner by using a non-homogeneous S-matrix algorithm (Pisarenco et al., 2011) which is an adaptation of the classical (homogeneous) S-matrix algorithm to the contrast field formulation.

4. Non-homogeneous S-matrix algorithm for repeating slices

In this section we develop a dedicated non-homogeneous S-matrix algorithm for the vertical problem. The repeating patterns in the direction of slicing suggest that this property can be exploited in order to achieve even higher efficiency in terms of computational costs.

The interface relation (47a) is unstable in the sense that its direct use for sequential elimination of field amplitudes leads to large round-off errors (Ko and Inkson, 1988; Li, 1996a; Pisarenco et al., 2011). To avoid instability, we use an S-matrix representation. This representation maps the incoming waves on an interface to outgoing waves from that interface. For instance, the S-matrix relation associated with layers $l$ and $l + 1$ is (see Figure 7)

$$
\begin{bmatrix}
    c_{l+1}^+ \\
    c_{l+1}^-
\end{bmatrix} =
\begin{bmatrix}
    S_{l+1}^{11} & S_{l+1}^{12} \\
    S_{l+1}^{21} & S_{l+1}^{22}
\end{bmatrix}
\begin{bmatrix}
    c_l^+ \\
    c_l^-
\end{bmatrix} +
\begin{bmatrix}
    f_{l+1}^1 \\
    f_{l+1}^2
\end{bmatrix}.
$$

The pair $(S_{l+1}, f_{l+1})$ describes the scattering and source properties of the interface. It can be determined from the T-matrix $T_{l+1}$ and the vector $g_{l+1}$.
\[
\begin{align*}
S_{11}^{l+1} &= (T_{11}^{l+1} - T_{12}^{l+1}(T_{22}^{l+1})^{-1}T_{21}^{l+1})X, 	ag{48a} \\
S_{12}^{l+1} &= T_{12}^{l+1}(T_{22}^{l+1})^{-1}X, 	ag{48b} \\
S_{21}^{l+1} &= -(T_{22}^{l+1})^{-1}T_{21}^{l+1}X, 	ag{48c} \\
S_{22}^{l+1} &= (T_{22}^{l+1})^{-1}X, 	ag{48d} \\
f_{1}^{l+1} &= g_{1}^{l+1} - T_{12}^{l+1}(T_{22}^{l+1})^{-1}g_{2}^{l+1}, 	ag{48e} \\
f_{2}^{l+1} &= -T_{22}^{l+1}g_{2}^{l+1}, 	ag{48f}
\end{align*}
\]

where
\[
\begin{pmatrix}
T_{11}^{l+1} & T_{12}^{l+1} \\
T_{21}^{l+1} & T_{22}^{l+1}
\end{pmatrix}
= \frac{1}{2}
\begin{pmatrix}
W_{l+1}^{-1} & V_{l+1}^{-1} \\
-W_{l+1}^{-1} & V_{l+1}^{-1}
\end{pmatrix}
\begin{pmatrix}
W_{l} & W_{l} \\
-V_{l} & -V_{l}
\end{pmatrix},
\]
\[
\begin{pmatrix}
g_{1}^{l+1} \\
g_{2}^{l+1}
\end{pmatrix}
= \frac{1}{2}
\begin{pmatrix}
W_{l+1}^{-1} & V_{l+1}^{-1} \\
-W_{l+1}^{-1} & V_{l+1}^{-1}
\end{pmatrix}
\left( g_{l}(h_{l+1}) - g_{l+1}(h_{l+1}) \right).
\]

Computation of matrices \(W_{l}\) and \(V_{l} = W_{l}Q_{l}\) requires diagonalization of \(C_{l}\) and \(D_{l}\) (Section 3.4), which is a relatively expensive computation. The time and memory requirements are estimated by
\[
\begin{align*}
T_{\text{diag}} &= O(M'N^3), \tag{50a} \\
M_{\text{diag}} &= O(M'N^2), \tag{50b}
\end{align*}
\]

where \(M'\) is the number of interfaces per period.

We now consider the problem of replacing two neighboring interfaces by one equivalent interface, depicted in Figure 8. Let the pair \((S_{a,b}, f_{a,b})\) describe the interface(s) joining layers \(a\) and \(b\), and the pair \((S_{b,c}, f_{b,c})\) describe the interface(s) joining layers \(b\) and \(c\), \((a < b < c, a, b, c \in \mathbb{N})\).

\[
\begin{pmatrix}
c_{a}^{+} \\
c_{a}^{-}
\end{pmatrix}
= \begin{pmatrix}
S_{a,b}^{11} & S_{a,b}^{12} \\
S_{a,b}^{21} & S_{a,b}^{22}
\end{pmatrix}
\begin{pmatrix}
c_{a}^{+} \\
c_{b}^{+}
\end{pmatrix}
+ \begin{pmatrix}
f_{a,b}^{1} \\
f_{a,b}^{2}
\end{pmatrix},
\]
\[
\begin{pmatrix}
c_{b}^{+} \\
c_{b}^{-}
\end{pmatrix}
= \begin{pmatrix}
S_{b,c}^{11} & S_{b,c}^{12} \\
S_{b,c}^{21} & S_{b,c}^{22}
\end{pmatrix}
\begin{pmatrix}
c_{b}^{+} \\
c_{c}^{+}
\end{pmatrix}
+ \begin{pmatrix}
f_{b,c}^{1} \\
f_{b,c}^{2}
\end{pmatrix}.
\]

We would like to find the pair \((S_{a,c}, f_{a,c})\) which realizes the mapping for layers \(a\) and \(c\)
\[
\begin{pmatrix}
c_{c}^{+} \\
c_{c}^{-}
\end{pmatrix}
= \begin{pmatrix}
S_{a,c}^{11} & S_{a,c}^{12} \\
S_{a,c}^{21} & S_{a,c}^{22}
\end{pmatrix}
\begin{pmatrix}
c_{a}^{+} \\
c_{c}^{+}
\end{pmatrix}
+ \begin{pmatrix}
f_{a,c}^{1} \\
f_{a,c}^{2}
\end{pmatrix}.
\]

(see (Pisarenco et al., 2011)),
It is shown in (Pisarenco et al., 2011) that \((S_{a,c}, f_{a,c})\) is given by
\[
S_{a,c} = \left( S_{b,c}^{11}(I - S_{a,b}^{21})^{-1} S_{a,b}^{11} + S_{b,c}^{12}(I - S_{a,b}^{21})^{-1} S_{a,b}^{12} \right) S_{b,c}^{22}(I - S_{a,b}^{21})^{-1} S_{a,b}^{22},
\]
\[
f_{a,c} = \left[ S_{b,c}^{11}(I - S_{a,b}^{21})^{-1} (S_{a,b}^{12} f_{b,c}^{1} + f_{b,c}^{1}) + S_{b,c}^{22}(I - S_{a,b}^{21})^{-1} (S_{a,b}^{12} f_{b,c}^{2} + f_{b,c}^{2}) \right].
\]

To simplify further exposition, based on the classical Redheffer matrix product (Redheffer, 1961), we define a non-homogeneous Redheffer product
\[
(S_{a,b}, f_{a,b}) \ast (S_{b,c}, f_{b,c}) = (S_{a,c}, f_{a,c}).
\]

It can be proven that this new product, just like the regular matrix product and the Redheffer matrix product, is associative but not commutative. We will exploit the associativity property in order to group the interfaces in a convenient way.

4.1. Global stack description

We can repeat the process of merging interfaces until the whole stack of interfaces is described by one equivalent interface. The non-homogeneous S-matrix algorithm (Pisarenco et al., 2011) uses the iteration
\[
(S_{1,l+1}, f_{1,l+1}) = (S_{1,l}, f_{1,l}) \ast (S_{l,l+1}, f_{l,l+1}).
\]

The global matrix-vector pair for \(M - 1\) interfaces (\(M\) layers) is given by
\[
(S_{1,M}, f_{1,M}) = [...[[ (S_{1,2}, f_{1,2}) \ast (S_{2,3}, f_{2,3}) ] \ast (S_{3,4}, f_{3,4}) ] \ast \ldots \ast (S_{M-1,M}, f_{M-1,M})].
\]
Figure 9: A schematic representation of the nonhomogeneous S-matrix algorithm applied to a grating with $R = 4$ periods and $M' = 2$ interfaces per period: classical linear recursion (left side) and fast exponential recursion (right side).

The left side of Figure 9 gives a visual representation of the merging process.

We consider a structure with $R$ repeating periods for simplicity assumed to be a power of 2, $R = 2^K$ (the algorithm can be extended to an arbitrary number of periods), where each period has $M'$ interfaces in the discretization ($M - 1 = RM'$).

The cost of computing (57) scales linearly with the number of periods,

\[
\mathcal{T}_S = \mathcal{O}(RM'N^3), \quad (58a)
\]
\[\mathcal{M}_S = \mathcal{O}(RM'N^2). \quad (58b)\]

Due to the associativity property of the non-homogeneous Redheffer product, we can regroup the multiplication operations. This allows us to exploit the local periodicity and reduce the computational costs down to

\[
\mathcal{T}_S = \mathcal{O}((M' + \log_2(R))N^3), \quad (59a)
\]
\[\mathcal{M}_S = \mathcal{O}((M' + \log_2(R))N^2). \quad (59b)\]
We observe that once we have computed the matrix-vector corresponding to the stack \( \{1, \ldots, M' + 1\} \), the matrix-vector corresponding to the stack \( \{M' + 1, \ldots, 2M' + 1\} \) can be computed with a minimal effort

\[
(S_{M' + 1, 2M' + 1}, f_{M' + 1, 2M' + 1}) = (S_{1, M' + 1}, \nu f_{1, M' + 1}),
\]

(60)

where \( \nu = \exp(-k_0 q(h_{M' + 1} - h_1)) \) is a phase factor which appears due to the scalar \( z \)-dependence of \( f \). Thus, we can directly compute the matrix-vector corresponding to the stack \( \{1, \ldots, 2M' + 1\} \)

\[
(S_{1, 2M' + 1}, f_{1, 2M' + 1}) = (S_{1, M' + 1}, f_{1, M' + 1}) \ast (S_{1, M' + 1}, \nu f_{1, M' + 1}).
\]

(61)

The iteration which computes the matrix-vector pair for an exponentially increasing number of layers is defined by

\[
(S_{1, 2^k + 1, M' + 1}, f_{1, 2^k + 1, M' + 1}) = (S_{1, 2^k + 1, M' + 1}, f_{1, 2^k + 1, M' + 1}) \ast (S_{1, 2^k + 1, M' + 1}, \nu^{2^k} f_{1, 2^k + 1, M' + 1}),
\]

(62)

for \( k = 0, \ldots, R - 1 \).

Once the matrix-vector pair for the entire stack is computed, the coefficients in the upper and lower layers are given by

\[
\begin{pmatrix}
  c_M^+ \\
  c_1^-
\end{pmatrix} =
\begin{pmatrix}
  f_{1,M}^1 \\
  f_{1,M}^2
\end{pmatrix}.
\]

(63)

The last expression is a result of (52) evaluated for \( a = 1 \) and \( c = M \).

4.2. Intermediary coefficients

Relation (63) allows the computation of the coefficients in the first and last layers of the stack. In order to compute the coefficients in the intermediary layers, we start with the matrix-vector pair associated with the whole stack and recursively halve the stack. One iteration of this process is explained by again considering the problem in Figure 8. This time we are given the pairs \( (S_{a,b}, f_{a,b}), (S_{b,c}, f_{b,c}) \) and the amplitudes \( c_a^+, c_c^- \), and we need to determine the intermediary coefficients \( c_b^+ \) and \( c_b^- \). Substitution of the second equation of (51b) in the first equation of (51a) yields

\[
(I - S_{a,b}^{12} S_{b,c}^{21}) c_b^+ = S_{a,b}^{11} c_a^+ + S_{a,b}^{12} S_{b,c}^{22} c_c^- + S_{a,b}^{12} f_{b,c}^2 + f_{a,b}^1.
\]

(64)
Substitution of the first equation of (51a) in the second equation of (51b) yields

\[
(I - S_{b,c}^{21} S_{a,b}^{12}) c_b^- = S_{b,c}^{21} S_{a,b}^{11} c_a^+ + S_{b,c}^{22} c_c^- + S_{b,c}^{21} f_{a,b}^1 + f_{b,c}^2.
\]

We can now write the amplitudes in layer \( b \) in terms of the amplitudes of the incoming modes in layer \( a \) and \( c \),

\[
\begin{pmatrix}
  c_b^+ \\
  c_b^-
\end{pmatrix} = \begin{pmatrix}
  (I - S_{a,b}^{12} S_{a,b}^{21})^{-1} S_{a,b}^{11} S_{a,b}^{12} (I - S_{b,c}^{21} S_{a,b}^{12})^{-1} S_{b,c}^{22} (I - S_{b,c}^{21} S_{a,b}^{12})^{-1} S_{b,c}^{21} f_{a,b}^1 + f_{b,c}^2 \\
  S_{b,c}^{21} f_{a,b}^1 + f_{b,c}^2
\end{pmatrix} \begin{pmatrix}
  c_a^+ \\
  c_c^-
\end{pmatrix}.
\]

The update relation (66) is applied recursively (by halving the stack) until the coefficients in all layers are determined. This requires \( RM' - 1 \) evaluations of (66) which, if computed in the right order, involves only matrix-vector products. Therefore, the computational cost of computing and storing the intermediary coefficients is given by

\[
T_c = O(RM'N^2),
\]

\[
M_c = O(RM'N^2).
\]

5. Summary of computational costs

From the point of view of computational cost the AFMM-CFF with classical and alternative discretizations consists of three consecutive tasks:

1. Matrix diagonalization (or eigenvalue decomposition);
2. Calculation of the global S-matrix;
3. Calculation of the intermediary coefficients.

The computational costs (in terms of time and memory) for each of the tasks is given respectively by (50), (59) and (67). By combining these estimates we obtain the total computational costs,

\[
T = ((c_T^{\text{diag}} + c_T^S) M' + c_T^S \log_2(R))N^3 + c_T^R RM'N^2,
\]

\[
M = ((c_M^{\text{diag}} + c_M^S) M' + c_M^S \log_2(R))N^2 + c_M^R RM'N.
\]

We specify that \( R \) in the above expressions gives the number of repeating patterns in the vertical direction. Thus, for horizontally repeating patterns
we always have $R = 1$ and $M'$ represents the total number of interfaces. With this convention estimate (68) is general and applicable for the vertical as well as the horizontal problem. For a reasonably large $N$, the second term in the above expressions (being one order lower) can be neglected. We will also assume (supported by evidence from practical experiments) that $c_{\text{diag}} = c_T = c_S$. Then estimate (68) reduces to

$$T \approx (2c_T M' + c_T \log_2(R)) N^3, \quad (69a)$$

$$M \approx ((c_{\text{diag}}^M + c_S^M) M' + c_S^M \log_2(R)) N^2. \quad (69b)$$

6. Results

We consider the problem of scattering from finite gratings. These are periodic structures with a finite number of periods. The problem is inspired from real-life applications in lithography, where gratings are printed on the wafer to be later used for quality control. By measuring the light scattered from a grating, its shape can be determined and the quality of the lithographic process can be assessed.

We first consider a small grating, infinitely long in the $y$-direction. The geometry of the problem is shown in Figure 10. It consists of $R = 8$ rectangular lines made of resist (with a refractive $n_{\text{resist}} = 1.5$), referred to as scatterer, supported by a silicon substrate ($n_{\text{silicon}} = 4.28 - 0.05i$). The material above the substrate and the grating lines is air ($n_{\text{air}} = 1.0$). The refractive indices $n$ of the material are related to the electric permittivity through

$$\epsilon(x, z) = n^2(x, z). \quad (70)$$

We obtain the numerical solution by using the classical and alternative discretization approaches for AFMM-CFF or, as explained in Section 2, by solving a horizontal problem and a vertical problem. For the horizontal problem, the incident plane wave is given by (3c) with

$$\tilde{a} = (0, 1, 0)^T,$$

$$\tilde{K}^{\text{inc}} = 10(\sqrt{2}/2, 0, \sqrt{2}/2)^T.$$

and the PMLs are placed in the regions $x \in [0, 0.1] \cup [1.8, 1.9]$. The distance units are $\mu m$. For the vertical problem, the amplitude and the wave vectors
of the incident plane wave are given by
\[
\bar{a} = (0, -1, 0)^T, \\
\bar{k}^{inc} = 10(\sqrt{2}/2, 0, \sqrt{2}/2)^T.
\]
and the PMLs are placed in the regions \(x \in [0, 0.1] \cup [0.4, 0.5]\). Both problems use a PML defined by (11) with \(p = 1\) and \(\beta_0 = 6\). It follows from the wavevectors \(\bar{k}^{inc}\) that in both cases the incident plane wave is TE-polarized and has a wavelength \(\lambda = 2\pi/|\bar{k}^{inc}| \approx 0.6283\mu m\). Note that the spatial discretization requires \(\bar{M} = 3\) slices in the horizontal problem and \(\bar{M} = \bar{M}' + 1 = 17\) slices in the vertical problem.

Figure 10 shows the solutions of the two problems. The solution of the vertical problem is plotted in rotated coordinates \((z, -y, x)\) to make the comparison with the horizontal problem easier. Clearly, for both problems the fields decay to zero in the PMLs. The solution in the PML is not physically relevant and the placement of the PMLs is different in the two problems. Therefore, the horizontal problem has a relevant solution in the stripe \((x, y, z) \in \bar{\Omega}_r = [0, 1.8] \times \mathbb{R} \times \mathbb{R}\), while the vertical problem has a relevant solution on \((z, -y, x) \in \bar{\Omega}_r = \mathbb{R} \times \mathbb{R} \times [1, 4]\). In the light of the above
Figure 11: Comparison of classical and alternative discretization approaches for a grating with 32 rectangular lines on resist substrate: (a) Convergence, (b) Complexity, (c) Error vs. Memory, (d) Error vs. Time. Solid line (—) corresponds to the horizontal problem (classical discretization), dashed line (— —) corresponds to the vertical problem (alternative discretization), dotted line (···) indicates theoretical cubic complexity, $T = O(N^3)$.

discussion, the fields should be compared only on a domain where both are relevant, that is on

$$\Omega_c = \bar{\Omega}_r \cup \bar{\Omega}_v = [0.1, 1.8] \times \mathbb{R} \times [0.1, 0.4]$$

Thus we expect that

$$\bar{e}^{c,N}(x) \approx T \bar{e}^{c,N}(T x), \text{ for } x = (x, y, z) \in \Omega_c$$  \hspace{1cm} (71)

A visual inspection of the obtained solutions in $\Omega_c$ confirms this expectation. Note that $N$ in (71) may be different for the horizontal and vertical problems.

In order to clearly assess the advantages of alternative discretization over classical discretization, we need to consider convergence and complexity of
the two approaches. We define first the discretization errors,

\[ \bar{\mathcal{E}}^N = ||\vec{e}^{c,N}(x) - \vec{e}^{c,ref}(x)|| \]  

\[ \bar{\mathcal{E}}^N = ||T\vec{e}^{c,N}(Tx) - \vec{e}^{c,ref}(x)|| \]  

where the norm for a vector field \( v \) is defined as

\[ ||v||^2 = \sum_i \sum_j \sum_{\alpha} |v_\alpha(x_i, 0, z_j)|^2, \]  

for \( \alpha = (x, y, z) \) and \( x_i \) and \( z_j \) sample the smallest rectangular domain enclosing the scatterer. For scatterers invariant in the \( y \)-direction, the dependence of the field on \( y \) is known, and only the field at \( y = 0 \) may be considered. In absence of an exact solution for this type of problems, we choose as reference the numerical solution for the (trusted) horizontal setup computed with a large truncation number \( N \). The computations are performed in MATLAB on a Intel Core 2 Quad CPU Q6600 4x2.40GHz with 2GB of RAM.

Figure 11 (a) shows the convergence of the solution for a vertical and horizontal setup. For small truncation numbers, \( N < 10 \), the error for the vertical problem is larger. The reason for this behavior lies in the representation of the vertical background field in the available basis. While in the horizontal problem a single basis function is required for an exact representation, in the vertical problem the background field \( e^b, h^b \) can only be approximated. Clearly, for small \( N \) the accuracy of the approximation is unsatisfactory. The slow decrease of the error for both approaches is explained by an insufficient resolution of the basis functions, so that the exponential decay of the solution in the PMLs cannot be accurately resolved. For \( N > 10 \), the error for the vertical problem \( \bar{\mathcal{E}} \) becomes much smaller than for the horizontal one \( \bar{\mathcal{E}} \). Because the width of the computational domain for \( \bar{P} \) is smaller than the width of the computational domain for \( P \), less basis functions are needed for the former to achieve the same accuracy. From the plot we conclude that the same number of harmonics determine an error for the vertical problem which is two orders of magnitude smaller than for the horizontal problem. However this plot is not sufficient to judge the performance of the approaches. Indeed, for a fixed number of harmonics, more work needs to be done for the vertical problem than for the horizontal due to a larger number of slices.
We turn to Figure 11 (b) which shows the execution time as a function of \( N \) for the two problems. Since \( \bar{M}' = 2, \bar{R} = 1 \) and \( \bar{M}' = 2, \bar{R} = 32 \), using (69) we can estimate the ratio of computation times for a large (but fixed) \( N \):

\[
\frac{(2c^T \bar{M}' + c^T \log_2(\bar{R})) \bar{N}^3}{(2c^T \bar{M}' + c^T \log_2(\bar{R})) \bar{N}^3} = \frac{4 + 0}{4 + 5} = \frac{1}{2.25}.
\]

Thus solving the vertical problem is estimated to take 2.25 times longer than the horizontal problem. The distance between the two lines on the plot for \( N \) approaching \( 10^3 \) comes close to this value. We also note that although asymptotically we expect a cubic complexity \( T = \mathcal{O}(N^3) \) indicated by the dotted line in the plot, it is only approached for \( N > 10^2 \). This means that the real speed-up factor will be lower than the one predicted theoretically.

We combine data from Figure 11 (a) and (b) in order to obtain the direct dependence of error on time displayed in Figure 11 (d). We conclude that for errors larger than \( 10^{-2} \) the vertical approach is slower. It only becomes faster for smaller errors. For instance, an error of \( 10^{-3} \) can be achieved approximately 8 times faster by solving \( \bar{P} \) instead of \( \bar{P} \).

For large computations with AFMM-CFF, the memory requirements limit the number of harmonics to be used. For this reason we also consider the memory required by the horizontal and vertical approaches to achieve a certain error. The size of a matrix or vector element in a double precision floating-point arithmetic is 8 Bytes. Considering that we have about 4 matrices per layer, the estimated memory space required for a computation is given by (68b) with \( c^M_{\text{diag}} = c^S = 36 \) and \( c^M_c = 16 \). Thus, the information from the convergence plot, Figure 11 (a), can be used to determine the error-vs-memory behavior displayed in Figure 11 (c).

We expect the gain in computational costs to change depending on the width of the structures. For this reason we now consider three geometries with the same material properties but having a different number of periods. Using the functions \( \bar{T}(\mathcal{E}) \) and \( \bar{\bar{T}}(\mathcal{E}) \) that give the computation time for the two problems, we define the speed-up factor for a given error

\[
\eta_T(\mathcal{E}) = \frac{\bar{T}(\mathcal{E})}{\bar{\bar{T}}(\mathcal{E})}.
\]

(75)

In analogy with the speed-up factor we also define the memory use factor

\[
\eta_M(\mathcal{E}) = \frac{\bar{M}(\mathcal{E})}{\bar{\bar{M}}(\mathcal{E})}.
\]

(76)
A plot of the speed-up and memory use factor for different values of $R$ ($R \in \{16, 32, 64\}$) is shown in Figure 12. We observe the general trend of increasing factors for decreasing errors. For $E = 10^{-3}$ the speed-up and memory use factors are in the range $3 \ldots 15$ and $10 \ldots 110$ respectively. Unfortunately it is difficult to simulate larger gratings ($R \in \{128, 256, \ldots\}$) with classical discretization (horizontal problem) due to the prohibitively large (in terms of memory) number of harmonics. This is also the reason why the error of $10^{-4}$ has not been reached for $R = 64$ (it is however easily reached in the vertical problem, see Figure 11). We can extrapolate the data presented in the plot in order to predict an approximate speed-up of $10^2$ and a memory use factor of $10^3$ at $E = 10^{-3}$ for a grating having 256 lines.

The convergence of the solution depends strongly on its smoothness, which in turn is determined by the presence and the size of the jump discontinuities in the discretized permittivity function $\epsilon_l(x)$. Clearly, these functions are different for the vertical and horizontal problems. The jump discontinuities are material interfaces located either at the boundary of the scatterer or at the boundaries of the background multilayer. Independent on the discretization direction, the discontinuities introduced by the scatterer have the same effect on the jumps in $\epsilon_l(x)$. The effect of the background multilayer however is different depending on the discretization direction. For the horizontal problem the jumps in $\bar{\epsilon}^b(z)$ are not visible in $\bar{\epsilon}_l(x)$ since they coincide with the interfaces of discretization slices. On the other hand, in the vertical problem the jump discontinuities in $\bar{\epsilon}^v(x)$ are present in $\bar{\epsilon}_l(x)$ for all layers.
This might slow down the convergence when solving the vertical problem. To quantify the effect, we consider a grating with $R = 32$ lines with three different materials in the substrate: air ($n_s = 1.0$), resist ($n_s = 1.5$), silicon ($n_s = 4.28 - 0.04i$). A plot of the speed-up and memory use factor for these cases is shown in Figure 13. In the case of air substrate the background permittivity function $\epsilon^b$ is continuous, which determines a faster convergence of the vertical problem’s solution. This results in large speed-up and memory use factors (up to 40 and 100 respectively). On the other hand, for silicon substrate we have a large jump discontinuity in $\epsilon^b$ at the interface between air and silicon and much smaller gains (2 and 8 for speed and memory respectively).

We can get a theoretical estimate of the attainable speed-up using (75) and (69)

$$\eta_T (\mathcal{E} \to 0) = \frac{2\bar{M} + \log_2(\bar{R})}{2\bar{M} + \log_2(\bar{R})} \left( \frac{\bar{N}}{N} \right)^3$$

We consider the case of a resist grating with 32 lines on air substrate studied in Figure 13. Because the superstrate and substrate are of the same material we have similar smoothness properties of the solution and permittivity in both directions. Then, we may assume that in order to achieve the same accuracy, we need to have equal resolution of the discretization. Thus, we require that the number of harmonics per unit length is constant for both...
discretizations. This yields the estimate

$$\tilde{N}/\bar{N} = \bar{\Lambda}/\bar{\Lambda} = 67/5 = 13.4.$$  \hfill (78)

We also have $\tilde{M}' = 2, \bar{R} = 1$ and $\tilde{M}' = 2, \bar{R} = 32$. Then, the speed-up factor estimated with (77) is close to 1000. We stress that this is a theoretical estimate for very large $N$ or, equivalently, for extremely small $E$. The estimate is useful as an upper bound for a potentially attainable speed-up. In practice, the pure cubic complexity in $N$ is hardly reached. For the considered problem, the error of $10^{-3}$ is reached for $N = 10^4..10^5$. In this range a quadratic complexity can be assumed, as confirmed by Figure 11 (b).

In this case our estimated speed-up factor becomes

$$\eta_T(E \approx 10^{-3}) = \frac{2\tilde{M}' + \log_2(\bar{R})}{2\bar{M}' + \log_2(\bar{R})} \left(\frac{\tilde{N}}{\bar{N}}\right)^2 = 79.8 \hfill (79)$$

which is reasonably close to the measured $\eta_T$ for $n_s = 1.0$ shown in Figure 13. It is important to stress that convergence depends on the smoothness of the permittivity. For problems which have different smoothness properties in different directions, the requirement on equal resolution (harmonics per unit length) in order to achieve equal errors (used in (78)) is not applicable. Moreover, in the vertical problem the background field has an exact representation only for a homogeneous background permittivity.

We mention that speed-up is achieved even if the structure has no repeating patterns. In this case the logarithmic complexity in the S-matrix algorithm is lost, but the cubic complexity can still be replaced by a linear complexity in the longer direction. We estimate the speed-up for an arbitrary structure with a width-to-height ratio given by $\bar{\Lambda}/\bar{\Lambda}$. We use expressions (77) and (79) with $\bar{R} = \bar{R} = 1$ (no repeating patterns), $\hat{M}'/\hat{M}' = \bar{\Lambda}/\bar{\Lambda}$ and $\hat{N}/\hat{N} = \bar{\Lambda}/\bar{\Lambda}$ (assuming similar smoothness of the permittivity in both directions) to get

$$\eta_T(E \to 0) = \left(\frac{\bar{\Lambda}}{\bar{\Lambda}}\right)^2, \hfill (80)$$

$$\eta_T(E \approx 10^{-3}) = \frac{\bar{\Lambda}}{\bar{\Lambda}}. \hfill (81)$$

Thus, the theoretical speed-up for an arbitrary structure scales quadratically with the width-to-height ratio. For accuracies required in practical applications this behavior will be closer to a linear one, as indicated by (81).
Although the most noticeable improvement resulting from the alternative discretization is speed-up, the memory saving might become crucial when solving large problems. Thus, with the new approach we may easily simulate scattering from structures with repeating patterns having a width of the order of hundreds of wavelengths. This is a difficult task for the AFMM-CFF with classical discretization, as well as for other numerical methods such as FDTD and FEM. Figure 14 shows a small part of the computed field for a grating with 1024 lines (more than 300 wavelengths).

In practical applications it is often needed to compute the field far above the scatterer. We briefly comment on this issue. The far field can be computed by a Green’s function approach (Michalski and Zheng, 1990) or by a Fourier transformation of the field on a line above the scatterer, denoted by a dotted line in Figure 2). The second approach (Goodman, 2004, Section 3.10.2) is faster since the integrals can be computed numerically with fast Fourier transform (FFT) routines. From Figure 2 it is visible that the field on the dotted line is not available in the PMLs and beyond for the horizontal problem. Therefore, the faster approach can only be used in the vertical problem.

7. Conclusions

A technique for speeding up computations with the AFMM-CFF has been presented. The speed-up is achieved by exchanging the directions of spatial and spectral discretizations. The computationally cheaper spatial discretization is used in the longer direction, while the more expensive spectral discretization is applied in the shorter direction. Moreover, local peri-
odicity in the slicing direction is exploited by using the associativity of the non-homogeneous Redheffer star product. This further decreases the computational costs: linear complexity is replaced by logarithmic complexity. The exchange of discretization directions is implemented by a rotation of the coordinate system. While the rotated scatterer can be treated automatically by the method, the rotated background multilayer introduces a fundamental difference: the background field cannot be exactly represented in the available basis and a projection has to be used. This step introduces an additional approximation error, especially when few basis functions are used. For this reason, larger speed-up is achieved when smaller errors are required (or equivalently more basis functions are employed). Besides the imposed accuracy, the speed-up and memory use factors obtained with alternative discretization are dependent on the number of repeating patterns as well as the size of jump discontinuities in the background multilayer. The numerical experiments confirm that the alternative discretization shows significant improvements for geometries with small jumps at the layer interfaces and a large number of periods.

In this paper we have considered two-dimensional scatterers (invariant in one direction). The presented approach can be extended to three-dimensional scatterers supported on multilayers invariant in the $x$- and $y$-directions. In this case instead of the classical approach (spectral discretization of $x$- and $y$-directions and spatial discretization of $z$) we can choose either $x$ or $y$ to be spatially discretized.

Acknowledgements

The authors acknowledge ASML B.V. for funding this project. We are also grateful to dr. Sjoerd Rienstra for his incisive question at a local seminar which has ignited this research.

References


### PREVIOUS PUBLICATIONS IN THIS SERIES:

<table>
<thead>
<tr>
<th>Number</th>
<th>Author(s)</th>
<th>Title</th>
<th>Month</th>
</tr>
</thead>
</table>
| 12-30  | M. Abdel Malik  
| 12-31  | F. Lippoth  
G. Prokert | Stability of equilibria for a two-phase osmosis model | Sept. ‘12 |
| 12-32  | G. Ali  
N. Banagaaya  
W.H.A. Schilders  
C. Tischendorf | Index-aware model order reduction for differential-algebraic equations | Oct. ‘12 |
| 12-33  | G. Ali  
N. Banagaaya  
W.H.A. Schilders  
| 12-34  | M. Pisarenco  
J.M.L. Maubach  
I.D. Setija  
R.M.M. Mattheij | Efficient solution of Maxwell’s equations for geometries with repeating patterns by an exchange of discretization directions in the aperiodic Fourier modal method | Oct. ‘12 |