Grating Diffraction Calculator (GD-Calc®) –

Coupled-Wave Theory for Biperiodic Diffraction Gratings

GD-Calc.pdf, version 09/17/2008

software.kjinnovation.com

1. Introduction

GD-Calc computes diffraction efficiencies of a biperiodic grating structure comprising linear, isotropic, and non-magnetic optical media. Part 1 of this document defines the class of optical geometries that can be modeled with GD-Calc, describes the electromagnetic field characteristics, and provides a conceptual framework for the software user interface. Part 2 describes the numerical algorithms used in the software code.

The grating diffraction theory developed in Part 2 is based on a generalization of the rigorous coupled-wave (RCW) method reviewed in Ref. 1. The general biperiodic grating theory has some commonality with the coupled-mode method described in Ref. 4 (e.g., the use of S-matrices [Ref. 2] and Fast Fourier Factorization [Ref. 3]); but the grating is described relative to a rectangular coordinate system in which only one of the grating period vectors need be aligned to the coordinate axes.

Part 1: User's Reference

2. Notation:

$(\ldots)$	grouping parentheses, matrices
$[\ldots]$	function arguments, superscripts
$\{\ldots\}$	set
$\hat{e}_1$, $\hat{e}_2$, $\hat{e}_3$	unit basis vectors
$\vec{x} = \hat{e}_1\, x_1 + \hat{e}_2\, x_2 + \hat{e}_3\, x_3$	position vector
$\vec{f} = \hat{e}_1\, f_1 + \hat{e}_2\, f_2 + \hat{e}_3\, f_3$	spatial frequency vector
$\vec{E} = \hat{e}_1\, E_1 + \hat{e}_2\, E_2 + \hat{e}_3\, E_3$	electric field vector
$\vec{H} = \hat{e}_1\, H_1 + \hat{e}_2\, H_2 + \hat{e}_3\, H_3$	magnetic field vector
ε	complex permittivity

3. Grating geometry

Figure 1 shows a biperiodic grating structure, which will be used to illustrate the types of grating geometries that can be modeled with GD-Calc. The example comprises square-section bars that are stacked to form an array of "#" structures. Only four such structures are shown, but the pattern extends periodically in two dimensions.

Position vectors are denoted $\vec{x} = \hat{e}_1\, x_1 + \hat{e}_2\, x_2 + \hat{e}_3\, x_3$, wherein $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ are orthonormal unit basis vectors. The grating structure is represented as a stack of grating strata bounded by planes of constant x_1, with boundaries at $x_1 = b_1[0], b_1[1], \ldots,$

$$b_1[l_1-1] < x_1 < b_1[l_1] \text{ in stratum } l_1;\ \ l_1 = 1 \ldots L_1 \tag{3.1}$$

wherein L_1 is the number of strata. (The "l_1" index will generally be used as a stratum identifier.) The strata do not necessarily represent physically distinct layers with different material compositions. Typically, the grating may be partitioned into strata in order to approximate sloped-wall layers by "staircase" profiles[1]. The grating is sandwiched between a substrate medium below the grating and a superstrate medium (e.g., vacuum) above the grating. In accordance with equation 3.1, the substrate and superstrate are considered to be semi-infinite strata defined by

$$-\infty = b_1[-1] < x_1 < b_1[0] \text{ in the substrate} \tag{3.2}$$

$$b_1[L_1] < x_1 < b_1[L_1+1] = +\infty \text{ in the superstrate} \tag{3.3}$$

Note that the x_1 coordinate increases toward the superstrate side of the grating and the strata are numbered from bottom (substrate) to top (superstrate). Incident illumination enters from the superstrate side.

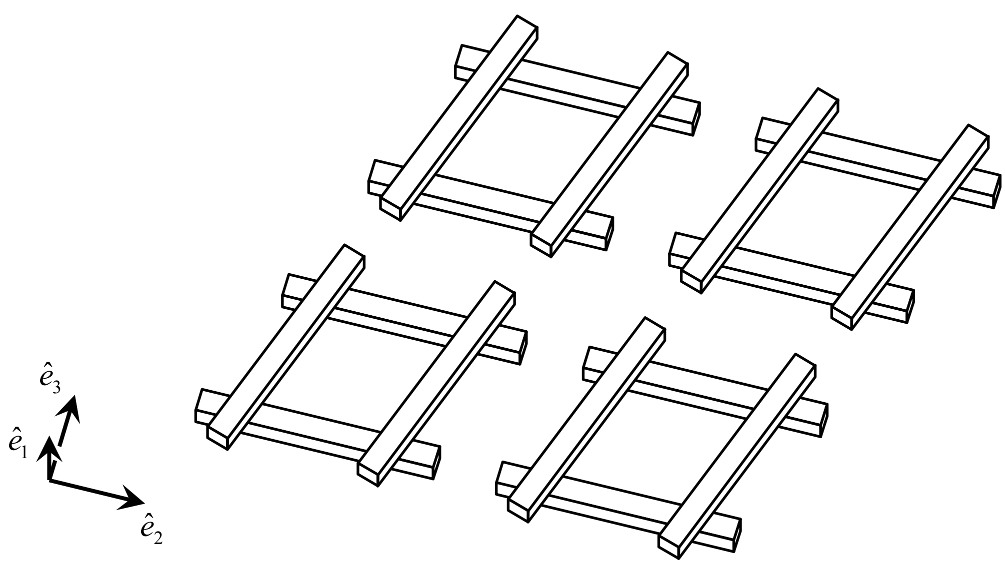

Figure 1. Biperiodic grating structure.

The grating is assumed to comprise isotropic, non-magnetic optical media, so its optical properties are fully characterized by its scalar, complex permittivity ε, which is a function of $\vec{x}$. This is the relative permittivity, equal to the square of the complex

[1] Stratification of sloped-wall profiles should be used with caution. Increasing the number of strata may not improve (and may actually degrade) calculation accuracy unless the number of diffraction orders is also significantly increased. This limitation of coupled-wave theory is discussed in Ref. 1, Sect. VI.5.

refractive index. ε is 1 in free space, and according to the assumed sign conventions the imaginary part of ε is non-negative,

$$\mathrm{Im}[\varepsilon] \geq 0 \tag{3.4}$$

The grating is characterized by two fundamental vector periods, $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$, which describe the permittivity function's translational symmetry characteristics. These vectors are parallel to the substrate and have the following coordinate representations,

$$\vec{d}_1^{[g]} = \hat{e}_2\, d_{2,1}^{[g]} + \hat{e}_3\, d_{3,1}^{[g]} \tag{3.5}$$

$$\vec{d}_2^{[g]} = \hat{e}_2\, d_{2,2}^{[g]} + \hat{e}_3\, d_{3,2}^{[g]} \tag{3.6}$$

The vectors are linearly independent,

$$d_{2,1}^{[g]}\, d_{3,2}^{[g]} \neq d_{3,1}^{[g]}\, d_{2,2}^{[g]} \tag{3.7}$$

The permittivity is invariant with respect to translational displacement by either vector $\vec{d}_1^{[g]}$ or $\vec{d}_2^{[g]}$,

$$\varepsilon[\vec{x} + \vec{d}_1^{[g]}] = \varepsilon[\vec{x}] \tag{3.8}$$

$$\varepsilon[\vec{x} + \vec{d}_2^{[g]}] = \varepsilon[\vec{x}] \tag{3.9}$$

For example, Figure 2 illustrates a plan view of the Figure 1 structure, showing the fundamental periods $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$.

$\varepsilon[\vec{x}]$ is constant outside of the grating structure, with a value $\varepsilon^{[\mathrm{sub}]}$ in the substrate and $\varepsilon^{[\mathrm{sup}]}$ in the superstrate, and it is independent of x_1 within each stratum. The top boundary coordinate for stratum l_1 (or for the substrate, if $l_1 = 0$) is denoted as $b_1[l_1]$,

$$\varepsilon[\vec{x}] = \varepsilon 1[l_1][x_2, x_3] \quad \text{for } b_1[l_1 - 1] < x_1 < b_1[l_1] \tag{3.10}$$

$$\varepsilon 1[0][x_2, x_3] = \varepsilon^{[\mathrm{sub}]} \tag{3.11}$$

$$\varepsilon 1[L_1 + 1][x_2, x_3] = \varepsilon^{[\mathrm{sup}]} \tag{3.12}$$

(The x_1 independence of $\varepsilon 1$ implies that the grating walls are perpendicular to the substrate within each stratum – hence the "staircase" approximation.) Based on equations 3.8 and 3.9, $\varepsilon 1$ satisfies the periodicity conditions

$$\varepsilon1[l_1][x_2 + d_{2,1}^{[g]}, x_3 + d_{3,1}^{[g]}] = \varepsilon1[l_1][x_2, x_3] \qquad (3.13)$$

$$\varepsilon1[l_1][x_2 + d_{2,2}^{[g]}, x_3 + d_{3,2}^{[g]}] = \varepsilon1[l_1][x_2, x_3] \qquad (3.14)$$

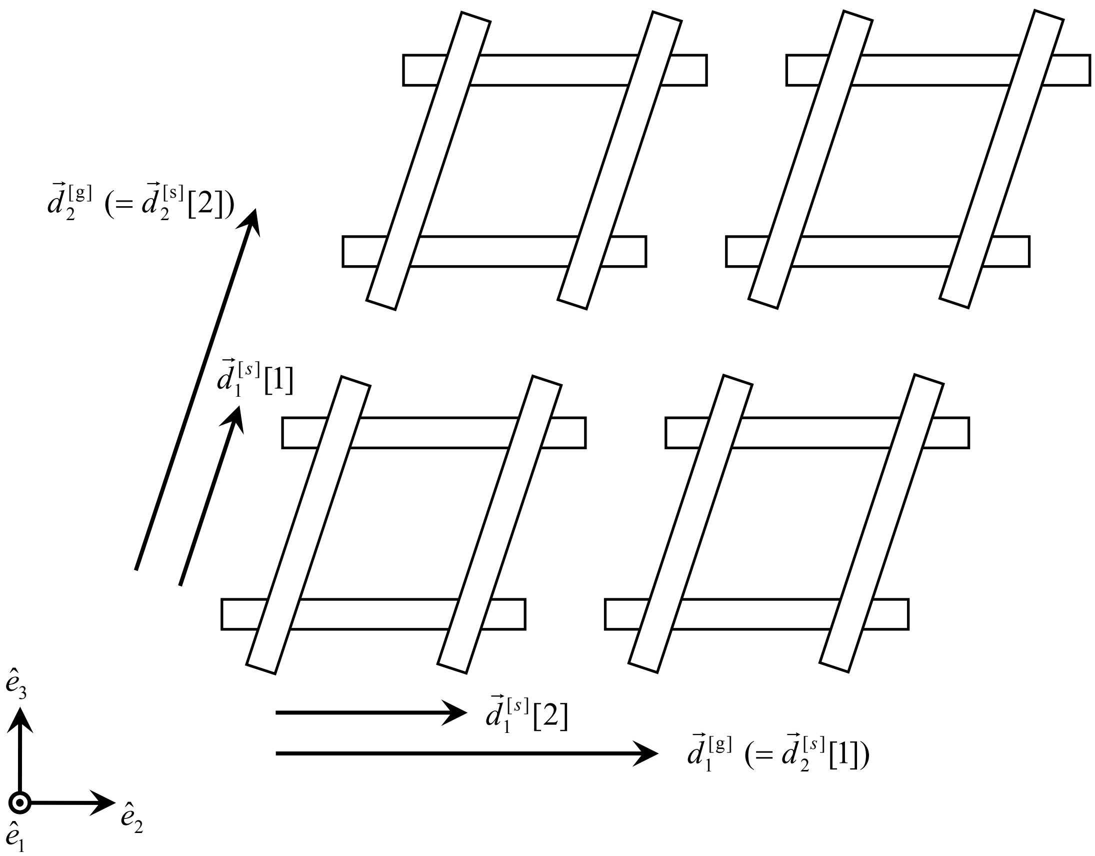

Figure 2. Grating periods.

Each stratum is characterized by at most two stratum-specific vector periods, $\vec{d}_1^{[s]}$ and $\vec{d}_2^{[s]}$ (not necessarily identical to $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$), which are parallel to the substrate and have coordinate representations similar to equations 3.5 and 3.6. The "s" superscript connotes "stratum", and the periods for stratum l_1 are denoted as $\vec{d}_1^{[s]}[l_1]$ and $\vec{d}_2^{[s]}[l_1]$. The permittivity in stratum l_1 is invariant with respect to translational displacement by either of these vectors,

$$\varepsilon1[l_1][x_2 + d_{2,1}^{[s]}[l_1], x_3 + d_{3,1}^{[s]}[l_1]] = \varepsilon1[l_1][x_2, x_3] \qquad (3.15)$$

$$\varepsilon1[l_1][x_2 + d_{2,2}^{[s]}[l_1], x_3 + d_{3,2}^{[s]}[l_1]] = \varepsilon1[l_1][x_2, x_3] \qquad (3.16)$$

These periodicity conditions may be stronger than conditions 3.13 and 3.14, which apply to all strata. For example, Figure 2 illustrates periods $\vec{d}_1^{[s]}[1]$ and $\vec{d}_2^{[s]}[1]$ for stratum 1, and $\vec{d}_1^{[s]}[2]$ and $\vec{d}_2^{[s]}[2]$ for stratum 2. $\vec{d}_1^{[s]}[l_1]$ and $\vec{d}_2^{[s]}[l_1]$ could be defined to be respectively equal to $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$ for all strata, but the geometry description will be simplified and the electromagnetic simulations may be more efficient if $\vec{d}_1^{[s]}[l_1]$ and $\vec{d}_2^{[s]}[l_1]$ are chosen so that their cross product, $\vec{d}_1^{[s]}[l_1] \times \vec{d}_2^{[s]}[l_1]$, has the smallest possible magnitude (i.e., the grating "unit cell" defined by $\vec{d}_1^{[s]}[l_1]$ and $\vec{d}_2^{[s]}[l_1]$ should preferably have minimal area).

The above description applies to the general case of a "biperiodic stratum" comprising a biperiodic grating structure. There are two special cases of strata that are treated differently than the biperiodic case.

First, a "uniperiodic stratum" comprises a lamellar line grating structure. This type of stratum is characterized by a single period $\vec{d}_1^{[s]}[l_1]$, which is (by convention) chosen to be perpendicular to the grating lines. (In this case, $\vec{d}_2^{[s]}[l_1]$ is implicitly perpendicular to $\vec{d}_1^{[s]}[l_1]$ and is of infinite length.)

Second, a "homogeneous stratum" comprises a homogeneous film, which is not characterized by periods. ($\vec{d}_1^{[s]}[l_1]$ and $\vec{d}_2^{[s]}[l_1]$ are both implicitly of infinite length in this case.)

Each stratum's type (biperiodic, uniperiodic, or homogeneous) and associated periods (if any) are defined in terms of the fundamental periods, $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$, and four stratum-specific "harmonic indices" $h_{1,1}[l_1]$, $h_{1,2}[l_1]$, $h_{2,1}[l_1]$, and $h_{2,2}[l_1]$ (corresponding to stratum l_1). The harmonic indices are integers, which define the stratum's periods as follows:

For a biperiodic stratum, the matrix of harmonic indices is non-singular,

$$h_{1,1}[l_1]\, h_{2,2}[l_1] \neq h_{1,2}[l_1]\, h_{2,1}[l_1] \quad \text{(biperiodic stratum)} \tag{3.17}$$

and the stratum's periods are defined by

$$\begin{pmatrix} d_{2,1}^{[s]}[l_1] & d_{2,2}^{[s]}[l_1] \\ d_{3,1}^{[s]}[l_1] & d_{3,2}^{[s]}[l_1] \end{pmatrix} = \begin{pmatrix} d_{2,1}^{[g]} & d_{2,2}^{[g]} \\ d_{3,1}^{[g]} & d_{3,2}^{[g]} \end{pmatrix} \begin{pmatrix} h_{1,1}[l_1] & h_{1,2}[l_1] \\ h_{2,1}[l_1] & h_{2,2}[l_1] \end{pmatrix}^{-1} \tag{3.18}$$

For example, in Figure 2 the following relations hold: $\vec{d}_1^{[g]} = \vec{d}_2^{[s]}[1] = 2\,\vec{d}_1^{[s]}[2]$, $\vec{d}_2^{[g]} = 2\,\vec{d}_1^{[s]}[1] = \vec{d}_2^{[s]}[2]$; or equivalently,

$$\begin{pmatrix} d_{2,1}^{[g]} & d_{2,2}^{[g]} \\ d_{3,1}^{[g]} & d_{3,2}^{[g]} \end{pmatrix} = \begin{pmatrix} d_{2,1}^{[s]}[1] & d_{2,2}^{[s]}[1] \\ d_{3,1}^{[s]}[1] & d_{3,2}^{[s]}[1] \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} d_{2,1}^{[s]}[2] & d_{2,2}^{[s]}[2] \\ d_{3,1}^{[s]}[2] & d_{3,2}^{[s]}[2] \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.19}$$

Hence, the harmonic coefficients for the two strata are

$$\begin{pmatrix} h_{1,1}[1] & h_{1,2}[1] \\ h_{2,1}[1] & h_{2,2}[1] \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} h_{1,1}[2] & h_{1,2}[2] \\ h_{2,1}[2] & h_{2,2}[2] \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.20}$$

For a uniperiodic stratum the harmonic indices satisfy the following conditions,

$$h_{1,1}[l1] \neq 0 \text{ or } h_{1,2}[l1] \neq 0; \quad h_{2,1}[l1] = 0 \text{ and } h_{2,2}[l1] = 0 \quad \text{(uniperiodic stratum)} \tag{3.21}$$

and the period $\vec{d}_1^{[s]}[l_1]$ is defined as follows: First compute the spatial frequency quantities

$$\left(f_{2,1}^{[s]}[l_1] \quad f_{3,1}^{[s]}[l_1] \right) = \left(h_{1,1}[l_1] \quad h_{1,2}[l_1] \right) \begin{pmatrix} d_{2,1}^{[g]} & d_{2,2}^{[g]} \\ d_{3,1}^{[g]} & d_{3,2}^{[g]} \end{pmatrix}^{-1} \tag{3.22}$$

and then define

$$\left(d_{2,1}^{[s]}[l_1] \quad d_{3,1}^{[s]}[l_1] \right) = \frac{\left(f_{2,1}^{[s]}[l_1] \quad f_{3,1}^{[s]}[l_1] \right)}{(f_{2,1}^{[s]}[l_1])^2 + (f_{3,1}^{[s]}[l_1])^2} \tag{3.23}$$

For a homogeneous stratum the harmonic indices are all zero,

$$h_{1,1}[l_1] = h_{1,2}[l_1] = h_{2,1}[l_1] = h_{2,2}[l_1] = 0 \quad \text{(homogeneous stratum)} \tag{3.24}$$

The above relations are conceptually simpler when expressed in terms of the grating's spatial frequencies. The grating has two fundamental spatial-frequency vectors $\vec{f}_1^{[g]}$ and $\vec{f}_2^{[g]}$, which have the coordinate representations

$$\vec{f}_1^{[g]} = \hat{e}_2\, f_{2,1}^{[g]} + \hat{e}_3\, f_{3,1}^{[g]} \tag{3.25}$$

$$\vec{f}_2^{[g]} = \hat{e}_2\, f_{2,2}^{[g]} + \hat{e}_3\, f_{3,2}^{[g]} \tag{3.26}$$

and which have the following reciprocal relationship to $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$,

$$\begin{pmatrix} f_{2,1}^{[\mathrm{g}]} & f_{3,1}^{[\mathrm{g}]} \\ f_{2,2}^{[\mathrm{g}]} & f_{3,2}^{[\mathrm{g}]} \end{pmatrix} \begin{pmatrix} d_{2,1}^{[\mathrm{g}]} & d_{2,2}^{[\mathrm{g}]} \\ d_{3,1}^{[\mathrm{g}]} & d_{3,2}^{[\mathrm{g}]} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.27}$$

Similarly, each stratum is characterized by its own spatial-frequency basis vectors $\vec{f}_1^{[s]}[l_1]$ and $\vec{f}_2^{[s]}[l_1]$ (for stratum l_1), which have the same reciprocal relationship to $\vec{d}_1^{[s]}[l_1]$ and $\vec{d}_2^{[s]}[l_1]$,

$$\vec{f}_1^{[s]}[l_1] = \hat{e}_2\, f_{2,1}^{[s]}[l_1] + \hat{e}_3\, f_{3,1}^{[s]}[l_1] \tag{3.28}$$

$$\vec{f}_2^{[s]}[l_1] = \hat{e}_2\, f_{2,2}^{[s]}[l_1] + \hat{e}_3\, f_{3,2}^{[s]}[l_1] \tag{3.29}$$

$$\begin{pmatrix} f_{2,1}^{[s]}[l_1] & f_{3,1}^{[s]}[l_1] \\ f_{2,2}^{[s]}[l_1] & f_{3,2}^{[s]}[l_1] \end{pmatrix} \begin{pmatrix} d_{2,1}^{[s]}[l_1] & d_{2,2}^{[s]}[l_1] \\ d_{3,1}^{[s]}[l_1] & d_{3,2}^{[s]}[l_1] \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.30}$$

For all three stratum types – biperiodic, uniperiodic, and homogeneous – a stratum's basis frequencies are a linear combination (i.e., "harmonics") of the grating's fundamental basis frequencies,

$$\begin{pmatrix} f_{2,1}^{[s]}[l_1] & f_{3,1}^{[s]}[l_1] \\ f_{2,2}^{[s]}[l_1] & f_{3,2}^{[s]}[l_1] \end{pmatrix} = \begin{pmatrix} h_{1,1}[l_1] & h_{1,2}[l_1] \\ h_{2,1}[l_1] & h_{2,2}[l_1] \end{pmatrix} \begin{pmatrix} f_{2,1}^{[\mathrm{g}]} & f_{3,1}^{[\mathrm{g}]} \\ f_{2,2}^{[\mathrm{g}]} & f_{3,2}^{[\mathrm{g}]} \end{pmatrix} \tag{3.31}$$

(By analogy with crystallography, the basis frequencies represent "reciprocal lattice vectors" and the harmonic indices are analogous to "Miller indices".) For a biperiodic stratum, $\vec{f}_1^{[s]}[l_1]$ and $\vec{f}_2^{[s]}[l_1]$ are non-zero and linearly independent; for a uniperiodic stratum, $\vec{f}_1^{[s]}[l_1]$ is non-zero and $\vec{f}_2^{[s]}[l_1]$ is zero; and for a homogeneous stratum $\vec{f}_1^{[s]}[l_1]$ and $\vec{f}_2^{[s]}[l_1]$ are both zero.

The grating structure within biperiodic stratum l_1 comprises rectangular "structural blocks" whose walls are parallel and perpendicular to $\vec{d}_2^{[s]}[l_1]$ (see Figure 2). The stratum geometry is defined by first partitioning the stratum into infinite-length "stripes" parallel to $\vec{d}_2^{[s]}[l_1]$, and then partitioning each stripe into blocks. For example, Figure 3 illustrates the partitioning of stratum 1 from the example of Figures 1 and 2. More complicated grating geometries can also be approximated as block-partitioned structures. For example, Figure 4 illustrates a possible representation for a circular structure[2].

[2] The preceding cautionary footnote regarding stratification also applies to block-partitioning of curved-wall structures such as circular posts and holes. Increasing the number of stripes may not improve calculation accuracy unless the number of diffraction orders is also significantly increased.

The stripes are numbered sequentially (e.g., “stripe 0”, “stripe 1”, “stripe 2”, etc.) in the order corresponding to the $\vec{d}_1^{[s]}[l_1]$ direction. The blocks within each stripe are numbered sequentially (e.g., “block 0”, “block 1”, etc.), with the block index order corresponding to the $\vec{d}_2^{[s]}[l_1]$ direction. Each block is optically homogeneous, and the permittivity in block l_3 of stripe l_2, stratum l_1 is denoted as $\varepsilon 3[l_1, l_2, l_3]$,

$$\varepsilon[\vec{x}] = \varepsilon 3[l_1, l_2, l_3] \quad \text{for } \vec{x} \text{ in block } l_3 \text{ of stripe } l_2\text{, stratum } l_1 \tag{3.32}$$

(The “l_2” index will generally be used as a stripe index, and “l_3” will be used as a structural block index.)

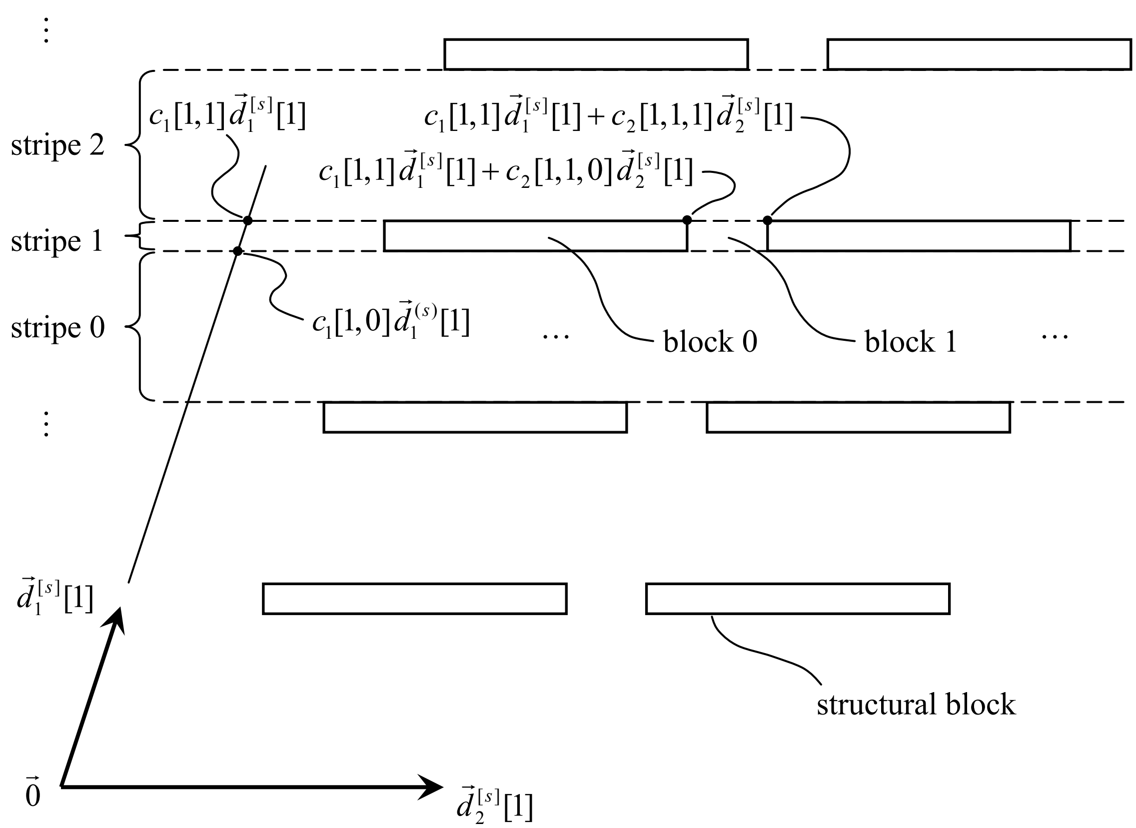

Figure 3. Stratum partitioning geometry.

The stripe boundary positions in stratum l_1 are defined in terms of a set of parameters $c_1[l_1, l_2]$, $l_2 = 1 \ldots L_2[l_1]$; wherein $L_2[l_1]$ is the number of stripes per period in the stratum. These values should satisfy the relation

$$c_1[l_1, L_2[l_1]] - 1 \le c_1[l_1, 1] \le c_1[l_1, 2] \le \ldots \le c_1[l_1, L_2[l_1]] \tag{3.33}$$

The l_2 range is implicitly extended to $\pm\infty$ by the condition

$$c_1[l_1, l_2 + L_2[l_1]] = c_1[l_1, l_2] + 1 \qquad (3.34)$$

(e.g., the left-hand value in relation 3.33 represents $c_1[l_1, 0]$). A ray from the coordinate origin ("$\vec{0}$" in Figure 3) in the $\vec{d}_1^{[s]}[l_1]$ direction intercepts the boundaries of stripe l_2 at points $c_1[l_1, l_2 - 1]\,\vec{d}_1^{[s]}[l_1]$ and $c_1[l_1, l_2]\,\vec{d}_1^{[s]}[l_1]$.

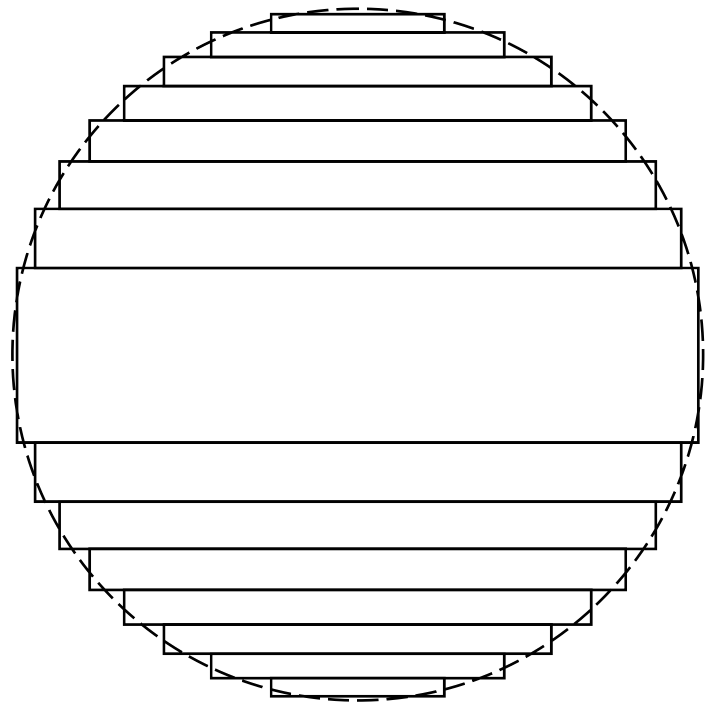

Figure 4. Block-partitioned circular structure.

The block boundary positions in stripe l_2 of stratum l_1 are defined in terms of a set of parameters $c_2[l_1, l_2, l_3]$, $l_2 = 1 \ldots L_2[l_1]$, $l_3 = 1 \ldots L_3[l_1, l_2]$; wherein $L_3[l_1, l_2]$ is the number of blocks per period in the stripe. These values should satisfy the relation

$$c_2[l_1, l_2, L_3[l_1, l_2]] - 1 \le c_2[l_1, l_2, 1] \le c_2[l_1, l_2, 2] \le \ldots \le c_2[l_1, l_2, L_3[l_1, l_2]] \qquad (3.35)$$

The l_2 and l_3 ranges are implicitly extended to $\pm\infty$ by the conditions

$$c_2[l_1, l_2, l_3 + L_3[l_1, l_2]] = c_2[l_1, l_2, l_3] + 1 \qquad (3.36)$$

$$c_2[l_1, l_2 + L_2[l_1], l_3] = c_2[l_1, l_2, l_3] \qquad (3.37)$$

$$L_3[l_1, l_2 + L_2[l_1]] = L_3[l_1, l_2] \tag{3.38}$$

(e.g. the left-hand value in relation 3.35 represents $c_2[l_1, l_2, 0]$).The boundaries of block l_3 in stripe l_2 intercept the interface between stripe l_2 and stripe $l_2 + 1$ at points $c_1[l_1, l_2]\, \vec{d}_1^{[s]}[l_1] + c_2[l_1, l_2, l_3 - 1]\, \vec{d}_2^{[s]}[l_1, l_2]$ and $c_1[l_1, l_2]\, \vec{d}_1^{[s]}[l_1] + c_2[l_1, l_2, l_3]\, \vec{d}_2^{[s]}[l_1, l_2]$.

A homogeneous stripe (e.g. stripe 0 in Figure 3) does not require any c_2 parameter specification; it only requires a c_1 parameter and a permittivity value. A uniperiodic stratum's stripes are all homogeneous and are aligned perpendicular to its $\vec{d}_1^{[s]}$ vector. A homogeneous stratum does not require any c_1 or c_2 specification; it requires only a permittivity specification.

A grating specification may contain include a coordinate break comprising lateral shift parameters Δx_2 and Δx_3, which implicitly translate strata laterally by the displacement vector

$$\Delta\vec{x} = \Delta x_2\, \hat{e}_2 + \Delta x_3\, \hat{e}_3 \tag{3.39}$$

The coordinate break is associated with a lateral plane at a particular x_1 height in the grating. The translational shift is only applied to strata above the coordinate break plane. Multiple coordinate breaks may be specified, and the total translation applied to any particular stratum is the sum of the $\Delta\vec{x}$ shifts specified by all the coordinate breaks below the stratum. (To apply a lateral translation to just a single stratum, a shift of $\Delta\vec{x}$ is applied immediately below the stratum and a shift of $-\Delta\vec{x}$ is applied immediately above the stratum.)

4. Electric field description

The electric field $\vec{E}$ comprises an incident field $\vec{E}^{[i]}$ and reflected field $\vec{E}^{[r]}$ in the superstrate, and a transmitted field $\vec{E}^{[t]}$ in the substrate. The incident field is assumed to be a single plane wave,

$$\vec{E}^{[i]}[\vec{x}] = \vec{A}^{[i]} \exp[i\, 2\pi\, \vec{f}^{[i]} \bullet \vec{x}] \tag{4.1}$$

wherein $\vec{A}^{[i]}$ is a constant vector and $\vec{f}^{[i]}$ is the incident field's spatial-frequency vector,

$$\vec{f}^{[i]} = \hat{e}_1\, f_1^{[i]} + \hat{e}_2\, f_2^{[i]} + \hat{e}_3\, f_3^{[i]} \tag{4.2}$$

The reflected field $\vec{E}^{[r]}$ consists of a superposition of plane-wave Fourier orders $\mathit{ff}\vec{E}^{[r]}[m_1, m_2]$ with spatial-frequency vectors $\vec{f}^{[r]}[m_1, m_2]$, which are labeled by two diffraction order indices m_1 and m_2,

$$\vec{E}^{[r]}[\vec{x}] = \sum_{m_1,\, m_2} \mathit{ff}\vec{E}^{[r]}[m_1, m_2][\vec{x}] \tag{4.3}$$

$$\mathit{ff}\vec{E}^{[r]}[m_1, m_2][\vec{x}] = \vec{A}^{[r]}[m_1, m_2]\exp[i\,2\pi\,\vec{f}^{[r]}[m_1, m_2] \bullet \vec{x}] \tag{4.4}$$

(The "f" prefix connotes a Fourier expansion, and "ff" connotes a two-dimensional Fourier expansion.) The transmitted field similarly consists of plane-wave Fourier orders $\mathit{ff}\vec{E}^{[t]}[m_1, m_2]$ with spatial-frequency vectors $\vec{f}^{[t]}[m_1, m_2]$

$$\vec{E}^{[t]}[\vec{x}] = \sum_{m_1,\, m_2} \mathit{ff}\vec{E}^{[t]}[m_1, m_2][\vec{x}] \tag{4.5}$$

$$\mathit{ff}\vec{E}^{[t]}[m_1, m_2][\vec{x}] = \vec{A}^{[t]}[m_1, m_2]\exp[i\,2\pi\,\vec{f}^{[t]}[m_1, m_2] \bullet \vec{x}] \tag{4.6}$$

The diffracted field's grating-tangential spatial frequencies (i.e., the $\hat{e}_2$ and $\hat{e}_3$ projections of $\vec{f}^{[r]}[m_1, m_2]$ and $\vec{f}^{[t]}[m_1, m_2]$) differ from that of the incident field by integer multiples of the grating's fundamental frequencies,

$$\begin{aligned}\left(f_2^{[r]}[m_1, m_2],\quad f_3^{[r]}[m_1, m_2]\right) &= \left(f_2^{[t]}[m_1, m_2],\quad f_3^{[t]}[m_1, m_2]\right)\\ &= \left(f_2^{[i]},\quad f_3^{[i]}\right) + m_1\left(f_{2,1}^{[\mathrm{g}]},\quad f_{3,1}^{[\mathrm{g}]}\right) + m_2\left(f_{2,2}^{[\mathrm{g}]},\quad f_{3,2}^{[\mathrm{g}]}\right)\end{aligned} \tag{4.7}$$

The plane waves' grating-normal spatial frequencies (the $\hat{e}_1$ frequency projections) are determined from the tangential frequencies,

$$f_1^{[i]} = -\sqrt{\frac{\varepsilon^{[\mathrm{sup}]}}{\lambda^2} - \left(f_2^{[i]}\right)^2 - \left(f_3^{[i]}\right)^2} \tag{4.8}$$

$$f_1^{[r]}[m_1, m_2] = +\sqrt{\frac{\varepsilon^{[\mathrm{sup}]}}{\lambda^2} - \left(f_2^{[r]}[m_1, m_2]\right)^2 - \left(f_3^{[r]}[m_1, m_2]\right)^2} \tag{4.9}$$

$$f_1^{[t]}[m_1, m_2] = -\sqrt{\frac{\varepsilon^{[\mathrm{sub}]}}{\lambda^2} - \left(f_2^{[t]}[m_1, m_2]\right)^2 - \left(f_3^{[t]}[m_1, m_2]\right)^2} \tag{4.10}$$

(The square root branch is chosen so that the square root's imaginary part is non-negative, and the square root signs in equations 4.8-10 are chosen so that the incident field propagates toward the grating and the diffracted fields propagate away from the grating.) Equations 4.7-10 define all of the field's spatial frequencies, based on a specification of $f_2^{[i]}$, $f_3^{[i]}$ and the grating frequencies.

In general, the order indices m_1 and m_2 range from $-\infty$ to $+\infty$, but for computational applications only a finite number of orders is retained. The retained orders are defined by the index limit sets $\mathcal{M}_1$ (which limits m_1) and $\mathcal{M}_2$ (which limits m_2),

$$m_2 \in \mathcal{M}_2 \tag{4.11}$$

$$m_1 \in \mathcal{M}_1[m_2] \tag{4.12}$$

The m_1 limit set $\mathcal{M}_1$ is a function of m_2. For example, Figure 5 illustrates a particular diffracted field's tangential frequencies (indicated as dots). The integer pairs represent diffraction order indices (m_1, m_2), and if only the labeled orders are intended to be retained then the index limit sets would be defined as follows,

$$\mathcal{M}_2 = \{-2, -1, 0, 1, 2\} \tag{4.13}$$

$$\left.\begin{array}{l} \mathcal{M}_1[-2] = \{0, 1, 2\} \\ \mathcal{M}_1[-1] = \{-1, 0, 1, 2\} \\ \mathcal{M}_1[\;0\;] = \{-2, -1, 0, 1, 2\} \\ \mathcal{M}_1[\;1\;] = \{-2, -1, 0, 1\} \\ \mathcal{M}_1[\;2\;] = \{-2, -1, 0\} \end{array}\right\} \tag{4.14}$$

The incident field amplitude is orthogonal to $\vec{f}^{[i]}$,

$$\vec{f}^{[i]} \bullet \vec{E}^{[i]} = 0 \tag{4.15}$$

$\vec{E}^{[i]}$ is specified in terms of its projections onto two unit vectors $\hat{s}^{[i]}$ and $\hat{p}^{[i]}$, wherein $\hat{s}^{[i]}$ is orthogonal to $\hat{e}_1$, and $\hat{s}^{[i]}$ and $\hat{p}^{[i]}$ are both orthogonal to $\vec{f}^{[i]}$,

$$\hat{s}^{[i]} = \hat{e}_2\, s_2^{[i]} + \hat{e}_3\, s_3^{[i]} \tag{4.16}$$

$$\hat{p}^{[i]} = \hat{e}_1\, p_1^{[i]} + \hat{e}_2\, p_2^{[i]} + \hat{e}_3\, p_3^{[i]} \tag{4.17}$$

$$\left(s_2^{[i]},\quad s_3^{[i]}\right)=\begin{cases}\dfrac{\left(-f_3^{[i]},\quad f_2^{[i]}\right)}{\sqrt{(f_2^{[i]})^2+(f_3^{[i]})^2}} & \text{if } (f_2^{[i]})^2+(f_3^{[i]})^2\neq 0\\ \dfrac{\left(-f_{3,1}^{[g]},\quad f_{2,1}^{[g]}\right)}{\sqrt{(f_{2,1}^{[g]})^2+(f_{3,1}^{[g]})^2}} & \text{if } (f_2^{[i]})^2+(f_3^{[i]})^2=0\end{cases}\tag{4.18}$$

$$\hat{p}^{[i]}=-\hat{s}^{[i]}\times\vec{f}^{[i]}\,\lambda/\sqrt{\varepsilon^{[\text{sup}]}}\tag{4.19}$$

The incident field is represented as

$$\vec{E}^{[i]}[\vec{x}]=\hat{s}^{[i]}\,E_s^{[i]}[\vec{x}]+\hat{p}^{[i]}\,E_p^{[i]}[\vec{x}]\tag{4.20}$$

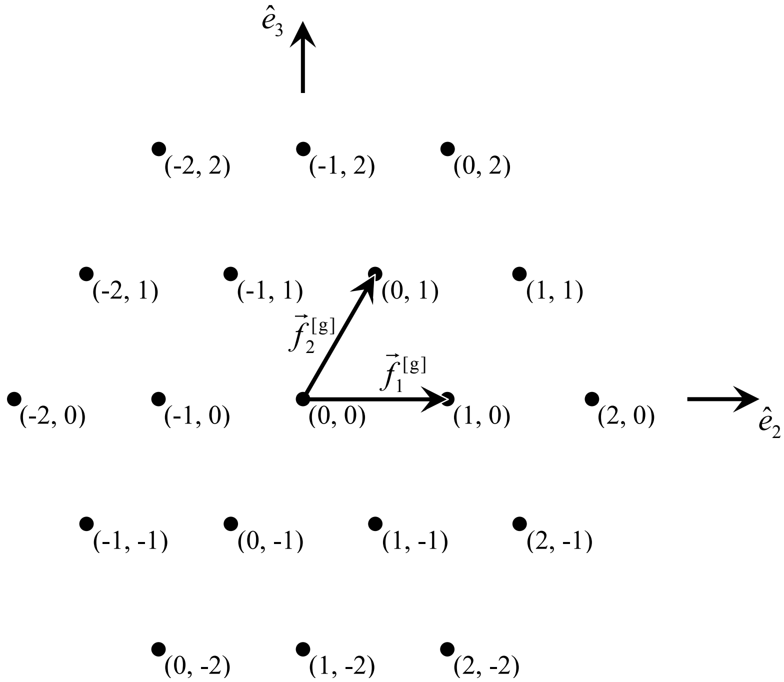

Figure 5. Electromagnetic field's tangential spatial frequencies.

$\hat{s}$ and $\hat{p}$ basis vectors are similarly defined for the reflected and transmitted orders, and the diffracted field amplitudes are projected onto these bases. For the reflected waves, the following definitions apply,

$$\hat{s}^{[r]}[m_1\,m_2]=\hat{e}_2\,s_2^{[r]}[m_1\,m_2]+\hat{e}_3\,s_3^{[r]}[m_1\,m_2]\tag{4.21}$$

$$\hat{p}^{[r]}[m_1\,m_2]=\hat{e}_1\,p_1^{[r]}[m_1\,m_2]+\hat{e}_2\,p_2^{[r]}[m_1\,m_2]+\hat{e}_3\,p_3^{[r]}[m_1\,m_2]\tag{4.22}$$

$$\left(s_2^{[r]}[m_1,m_2],\quad s_3^{[r]}[m_1,m_2]\right)=$$
$$\begin{cases}\dfrac{\left(-f_3^{[r]}[m_1,m_2],\quad f_2^{[r]}[m_1,m_2]\right)}{\sqrt{(f_2^{[r]}[m_1,m_2])^2+(f_3^{[r]}[m_1,m_2])^2}} & \text{if } (f_2^{[r]}[m_1,m_2])^2+(f_3^{[r]}[m_1,m_2])^2\neq 0\\ \dfrac{\left(-f_{3,1}^{[g]},\quad f_{2,1}^{[g]}\right)}{\sqrt{(f_{2,1}^{[g]})^2+(f_{3,1}^{[g]})^2}} & \text{if } (f_2^{[r]}[m_1,m_2])^2+(f_3^{[r]}[m_1,m_2])^2=0\end{cases}$$
(4.23)

$$\hat{p}^{[r]}[m_1,m_2]=+\hat{s}^{[r]}[m_1,m_2]\times\vec{f}^{[r]}[m_1,m_2]\,\lambda/\sqrt{\varepsilon^{[\text{sup}]}} \tag{4.24}$$

$$\begin{aligned}&f\!f\vec{E}^{[r]}[m_1,m_2][\vec{x}]=\\&\quad\hat{s}^{[r]}[m_1,m_2]\,f\!fE_s^{[r]}[m_1,m_2][\vec{x}]+\hat{p}^{[r]}[m_1,m_2]\,f\!fE_p^{[r]}[m_1,m_2][\vec{x}]\end{aligned} \tag{4.25}$$

For the transmitted waves, the definitions are

$$\hat{s}^{[t]}[m_1\,m_2]=\hat{e}_2\,s_2^{[t]}[m_1\,m_2]+\hat{e}_3\,s_3^{[t]}[m_1\,m_2] \tag{4.26}$$

$$\hat{p}^{[t]}[m_1\,m_2]=\hat{e}_1\,p_1^{[t]}[m_1\,m_2]+\hat{e}_2\,p_2^{[t]}[m_1\,m_2]+\hat{e}_3\,p_3^{[t]}[m_1\,m_2] \tag{4.27}$$

$$\left(s_2^{[t]}[m_1,m_2],\quad s_3^{[t]}[m_1,m_2]\right)=\left(s_2^{[r]}[m_1,m_2],\quad s_3^{[r]}[m_1,m_2]\right) \tag{4.28}$$

$$\hat{p}^{[t]}[m_1,m_2]=-\hat{s}^{[t]}[m_1,m_2]\times\vec{f}^{[t]}[m_1,m_2]\,\lambda/\sqrt{\varepsilon^{[\text{sub}]}} \tag{4.29}$$

$$\begin{aligned}&f\!f\vec{E}^{[t]}[m_1,m_2][\vec{x}]=\\&\quad\hat{s}^{[t]}[m_1,m_2]\,f\!fE_s^{[t]}[m_1,m_2][\vec{x}]+\hat{p}^{[t]}[m_1,m_2]\,f\!fE_p^{[t]}[m_1,m_2][\vec{x}]\end{aligned} \tag{4.30}$$

$\hat{s}^{[t]}$ is equal to $\hat{s}^{[r]}$ (equation 4.28) because all the terms on the right side of equation 4.23 would be the same with $\vec{f}^{[t]}$ substituted for $\vec{f}^{[r]}$ (equation 4.7). Note the sign convention for $\hat{p}$: Equations 4.19 and 4.29 include a minus sign, whereas equation 4.24 does not.

The diffracted fields are linearly dependent on the incident field, and the linear coefficients for order-(m_1,m_2) reflected and transmitted waves will be represented by reflection and transmission matrices $R[m_1,m_2]$ and $T[m_1,m_2]$,

$$R[m_1,m_2]=\begin{pmatrix}R_{s,s}[m_1,m_2] & R_{s,p}[m_1,m_2]\\ R_{p,s}[m_1,m_2] & R_{p,p}[m_1,m_2]\end{pmatrix} \tag{4.31}$$

$$T[m_1,m_2]=\begin{pmatrix} T_{s,s}[m_1,m_2] & T_{s,p}[m_1,m_2] \\ T_{p,s}[m_1,m_2] & T_{p,p}[m_1,m_2] \end{pmatrix} \tag{4.32}$$

In defining these matrices, the incident and reflected amplitudes are evaluated at point $\vec{x}=\hat{e}_1\, b_1[L_1]$ on the grating's top surface (equation 3.3), and the transmitted amplitudes are evaluated at point $\vec{x}=\hat{e}_1\, b_1[0]$ on the grating's bottom surface (equation 3.2),

$$\begin{pmatrix} ffE_s^{[r]}[m_1,m_2][\hat{e}_1\, b_1[L_1]] \\ ffE_p^{[r]}[m_1,m_2][\hat{e}_1\, b_1[L_1]] \end{pmatrix} = \begin{pmatrix} R_{s,s}[m_1,m_2] & R_{s,p}[m_1,m_2] \\ R_{p,s}[m_1,m_2] & R_{p,p}[m_1,m_2] \end{pmatrix} \begin{pmatrix} E_s^{[i]}[\hat{e}_1\, b_1[L_1]] \\ E_p^{[i]}[\hat{e}_1\, b_1[L_1]] \end{pmatrix} \tag{4.33}$$

$$\begin{pmatrix} ffE_s^{[t]}[m_1,m_2][\hat{e}_1\, b_1[0]] \\ ffE_p^{[t]}[m_1,m_2][\hat{e}_1\, b_1[0]] \end{pmatrix} = \begin{pmatrix} T_{s,s}[m_1,m_2] & T_{s,p}[m_1,m_2] \\ T_{p,s}[m_1,m_2] & T_{p,p}[m_1,m_2] \end{pmatrix} \begin{pmatrix} E_s^{[i]}[\hat{e}_1\, b_1[L_1]] \\ E_p^{[i]}[\hat{e}_1\, b_1[L_1]] \end{pmatrix} \tag{4.34}$$

The incident power $P^{[i]}$ propagating toward the grating (i.e., in the $-\hat{e}_1$ direction) is (within a dimensional constant)

$$P^{[i]} = -\left(|\, E_s^{[i]}\,|^2 + |\, E_p^{[i]}\,|^2\right)\mathrm{Re}[\lambda\, f_1^{[i]}] \tag{4.35}$$

Similarly, the reflected power $P^{[r]}[m_1,m_2]$ propagating away from the grating (in the $+\hat{e}_1$ direction) in the order-(m_1,m_2) reflected wave is

$$P^{[r]}[m_1,m_2] = \left(|\, ffE_s^{[r]}[m_1,m_2]\,|^2 + |\, ffE_p^{[r]}[m_1,m_2]\,|^2\right)\mathrm{Re}[\lambda\, f_1^{[r]}[m_1,m_2]] \tag{4.36}$$

and the transmitted power $P^{[t]}[m_1,m_2]$ propagating away from the grating (in the $-\hat{e}_1$ direction) in the order-(m_1,m_2) transmitted wave is

$$P^{[t]}[m_1,m_2] = -\left(|\, ffE_s^{[t]}[m_1,m_2]\,|^2 + |\, ffE_p^{[t]}[m_1,m_2]\,|^2\right)\mathrm{Re}[\lambda\, f_1^{[t]}[m_1,m_2]] \tag{4.37}$$

(The incident and reflected field amplitudes are implicitly evaluated at $\vec{x}=\hat{e}_1\, b_1[L_1]$ in equations 4.35 and 4.36, and the transmitted amplitudes are evaluated at $\vec{x}=\hat{e}_1\, b_1[0]$ in equation 4.37.) The order-(m_1,m_2) reflected and transmitted diffraction efficiencies $\eta^{[r]}[m_1,m_2]$ and $\eta^{[t]}[m_1,m_2]$ (i.e., ratio of diffracted to incident power for each diffracted order) are

$$\eta^{[r]}[m_1,m_2] = \frac{P^{[r]}[m_1,m_2]}{P^{[i]}} \\ = -\left(\frac{|\,\mathit{ff}E_s^{[r]}[m_1,m_2]\,|^2 + |\,\mathit{ff}E_p^{[r]}[m_1,m_2]\,|^2}{|\,E_s^{[i]}\,|^2 + |\,E_p^{[i]}\,|^2} \right) \frac{\mathrm{Re}[f_1^{[r]}[m_1,m_2]]}{\mathrm{Re}[f_1^{[i]}]} \tag{4.38}$$

$$\eta^{[t]}[m_1,m_2] = \frac{P^{[t]}[m_1,m_2]}{P^{[i]}} \\ = \left(\frac{|\,\mathit{ff}E_s^{[t]}[m_1,m_2]\,|^2 + |\,\mathit{ff}E_p^{[t]}[m_1,m_2]\,|^2}{|\,E_s^{[i]}\,|^2 + |\,E_p^{[i]}\,|^2} \right) \frac{\mathrm{Re}[f_1^{[t]}[m_1,m_2]]}{\mathrm{Re}[f_1^{[i]}]} \tag{4.39}$$

(For an evanescent wave in a lossless medium, f_1 is pure imaginary, so its diffraction efficiency is zero.)

Part 2: Theory and Methods

5. Fourier expansion of the electromagnetic field

The total electromagnetic field consists of the incident field (which has the form described by equation 4.1 for $\vec{E}$, and a similar form for $\vec{H}$), and the diffracted field (which includes the reflected and transmitted fields as well as the field within the grating). Since the grating is invariant under translation by period $\vec{d}_1^{[g]}$ (equation 3.8), it can be expected that a translation of the incident field by $\vec{d}_1^{[g]}$ will have no effect on the resulting diffracted field other than to translate it by the same offset. A translation of the incident field by $\vec{d}_1^{[g]}$ has the same effect as multiplying the field by a constant factor of $\exp[i\,2\pi\,\vec{f}^{[i]} \bullet \vec{d}_1^{[g]}]$,

$$\vec{E}^{[i]}[\vec{x} + \vec{d}_1^{[g]}] = \vec{E}^{[i]}[\vec{x}]\exp[i\,2\pi\,\vec{f}^{[i]} \bullet \vec{d}_1^{[g]}] \tag{5.1}$$

(from equation 4.1). The diffracted field can be expected to have a linear dependence on the incident field, so the resulting diffracted field will also be scaled by the same factor. Hence, a translation of the diffracted field by $\vec{d}_1^{[g]}$ is similarly equivalent to applying the above scaling factor, and therefore the above relation applies to the total (incident plus diffracted) field,

$$\vec{E}[\vec{x} + \vec{d}_1^{[g]}] = \vec{E}[\vec{x}]\exp[i\,2\pi\,\vec{f}^{[i]} \bullet \vec{d}_1^{[g]}] \tag{5.2}$$

Since $\vec{d}_1^{[g]}$ has no $\hat{e}_1$ component (equation 3.5), the scale factor in equation 5.2 only depends on the projection of $\vec{f}^{[i]}$ parallel to the grating surface, denoted as $\vec{f}^{[ip]}$,

$$\vec{f}^{[ip]} = \hat{e}_2 \, f_2^{[i]} + \hat{e}_3 \, f_3^{[i]} \tag{5.3}$$

$$\vec{E}[\vec{x} + \vec{d}_1^{[g]}] = \vec{E}[\vec{x}] \exp[i \, 2\pi \, \vec{f}^{[ip]} \bullet \vec{d}_1^{[g]}] \tag{5.4}$$

(The "p" in the superscript connotes "parallel".) This implies that the total field divided by $\exp[i \, 2\pi \, \vec{f}^{[ip]} \bullet \vec{x}]$ is periodic with period $\vec{d}_1^{[g]}$,

$$\vec{E}[\vec{x} + \vec{d}_1^{[g]}] / \exp[i \, 2\pi \, \vec{f}^{[ip]} \bullet (\vec{x} + \vec{d}_1^{[g]})] = \vec{E}[\vec{x}] / \exp[i \, 2\pi \, \vec{f}^{[ip]} \bullet \vec{x}] \tag{5.5}$$

The same type of condition also holds for the grating's second fundamental period, $\vec{d}_2^{[g]}$,

$$\vec{E}[\vec{x} + \vec{d}_2^{[g]}] / \exp[i \, 2\pi \, \vec{f}^{[ip]} \bullet (\vec{x} + \vec{d}_2^{[g]})] = \vec{E}[\vec{x}] / \exp[i \, 2\pi \, \vec{f}^{[ip]} \bullet \vec{x}] \tag{5.6}$$

The displacement periods $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$ in equations 5.5 and 5.6 are in the $\hat{e}_2$, $\hat{e}_3$ plane (equations 3.5, 3.6), so the right-hand expression in these equations can be represented as a biperiodic Fourier series in x_2 and x_3. Denoting the order-(m_1, m_2) Fourier coefficient as $ff\vec{E}[m_1, m_2][x_1]$, the following $\vec{E}$-field Fourier expansion is obtained,

$$\vec{E}[\vec{x}] = \sum_{m_1, m_2} ff\vec{E}[m_1, m_2][x_1] \exp[i \, 2\pi \, \vec{f}^{[p]}[m_1, m_2] \bullet \vec{x}] \tag{5.7}$$

wherein the field's grating-tangential spatial frequencies $\vec{f}^{[p]}[m_1, m_2]$ are

$$\vec{f}^{[p]}[m_1, m_2] = \vec{f}^{[ip]} + m_1 \, \vec{f}_1^{[g]} + m_2 \, \vec{f}_2^{[g]} \tag{5.8}$$

(The fundamental grating frequencies $\vec{f}_1^{[g]}$ and $\vec{f}_2^{[g]}$ are defined by equations 3.25-27, and the periodicity conditions 5.5 and 5.6 can be verified directly using equations 5.7 and 3.27.) The coordinate representation of $\vec{f}^{[p]}[m_1, m_2]$ is

$$\vec{f}^{[p]}[m_1, m_2] = \hat{e}_2 \, f_2^{[p]}[m_1, m_2] + \hat{e}_3 \, f_3^{[p]}[m_1, m_2] \tag{5.9}$$

$$f_2^{[p]}[m_1, m_2] = f_2^{[i]} + m_1 \, f_{2,1}^{[g]} + m_2 \, f_{2,2}^{[g]} \tag{5.10}$$

$$f_3^{[p]}[m_1, m_2] = f_3^{[i]} + m_1 \, f_{3,1}^{[g]} + m_2 \, f_{3,2}^{[g]} \tag{5.11}$$

The magnetic field $\vec{H}$ is represented by a similar Fourier expansion,

$$\vec{H}[\vec{x}] = \sum_{m_1, m_2} \mathit{ff}\vec{H}[m_1, m_2][x_1] \exp[i\, 2\pi\, \vec{f}^{[p]}[m_1, m_2] \bullet \vec{x}] \tag{5.12}$$

Based on the order limit conditions 4.11 and 4.12, the field expansions 5.7 and 5.12 are truncated as follows,

$$\vec{E}[\vec{x}] \cong \sum_{m_2 \in \mathcal{M}_2} \; \sum_{m_1 \in \mathcal{M}_1[m_2]} \mathit{ff}\vec{E}[m_1, m_2][x_1] \exp[i\, 2\pi\, \vec{f}^{[p]}[m_1, m_2] \bullet \vec{x}] \tag{5.13}$$

$$\vec{H}[\vec{x}] \cong \sum_{m_2 \in \mathcal{M}_2} \; \sum_{m_1 \in \mathcal{M}_1[m_2]} \mathit{ff}\vec{H}[m_1, m_2][x_1] \exp[i\, 2\pi\, \vec{f}^{[p]}[m_1, m_2] \bullet \vec{x}] \tag{5.14}$$

6. The Maxwell Equations and homogeneous-medium solutions

The Maxwell Equations (in Gaussian units) for a time-periodic electromagnetic field in a linear, isotropic, non-magnetic medium are

$$\nabla \times \vec{E} = i \tfrac{2\pi}{\lambda} \vec{H} \tag{6.1}$$

$$\nabla \times \vec{H} = -i \tfrac{2\pi}{\lambda} \varepsilon\, \vec{E} \tag{6.2}$$

wherein λ is the wavelength in vacuum and ε is the complex permittivity. An implicit time-separable factor of $\exp[-i\, 2\pi\, c\, t / \lambda]$ is assumed. (t is time and c is the speed of light in vacuum.) The minus sign in this time factor implies $\mathrm{Im}[\varepsilon] \geq 0$ (relation 3.4).

Considering the case where ε is equal to a constant $\varepsilon^{[\mathrm{c}]}$ (within some x_1 range),

$$\varepsilon[\vec{x}] = \varepsilon^{[\mathrm{c}]} \tag{6.3}$$

substitution from equations 5.7 and 5.12 in 6.1 and 6.2 yields

$$\begin{aligned} &i\, 2\pi\, \vec{f}^{[p]}[m_1, m_2] \times \mathit{ff}\vec{E}[m_1, m_2] + \hat{e}_1 \times \partial_1\, \mathit{ff}\vec{E}[m_1, m_2] = \\ &\quad i \tfrac{2\pi}{\lambda}\, \mathit{ff}\vec{H}[m_1, m_2] \end{aligned} \tag{6.4}$$

$$\begin{aligned} &i\, 2\pi\, \vec{f}^{[p]}[m_1, m_2] \times \mathit{ff}\vec{H}[m_1, m_2] + \hat{e}_1 \times \partial_1\, \mathit{ff}\vec{H}[m_1, m_2] = \\ &\quad -i \tfrac{2\pi}{\lambda} \varepsilon^{[\mathrm{c}]}\, \mathit{ff}\vec{E}[m_1, m_2] \end{aligned} \tag{6.5}$$

(∂_1 represents the derivative with respect to x_1.) Particular solutions of these equations are of the form

$$ff\vec{E}[m_1,m_2][x_1] = \vec{A}^{[E]} \exp[i\,2\pi\, f_1\, x_1] \tag{6.6}$$

$$ff\vec{H}[m_1,m_2][x_1] = \vec{A}^{[H]} \exp[i\,2\pi\, f_1\, x_1] \tag{6.7}$$

wherein $\vec{A}^{[E]}$, $\vec{A}^{[H]}$, and f_1 are undetermined constants. Defining

$$\vec{f} = \hat{e}_1\, f_1 + \hat{e}_2\, f_2 + \hat{e}_3\, f_3 \tag{6.8}$$

with

$$f_2 = f_2^{[p]}[m_1,m_2] \tag{6.9}$$

$$f_3 = f_3^{[p]}[m_1,m_2] \tag{6.10}$$

equations 6.4 and 6.5 reduce to

$$\vec{f} \times \vec{A}^{[E]} = \tfrac{1}{\lambda} \vec{A}^{[H]} \tag{6.11}$$

$$\vec{f} \times \vec{A}^{[H]} = -\tfrac{1}{\lambda} \varepsilon^{[\mathrm{c}]}\, \vec{A}^{[E]} \tag{6.12}$$

The unit vector $\hat{s}$ is defined by[3]

$$\hat{s} = \hat{e}_2\, s_2 + \hat{e}_3\, s_3 \tag{6.13}$$

$$(s_2,\quad s_3) = \begin{cases} \dfrac{(-f_3,\quad f_2)}{\sqrt{(f_2)^2 + (f_3)^2}} & \text{if } (f_2)^2 + (f_3)^2 \neq 0 \\ \dfrac{(-f_{3,1}^{[g]},\quad f_{2,1}^{[g]})}{\sqrt{(f_{2,1}^{[g]})^2 + (f_{3,1}^{[g]})^2}} & \text{if } (f_2)^2 + (f_3)^2 = 0 \end{cases} \tag{6.14}$$

$\vec{f}$, $\hat{s}$, and $\hat{s} \times \vec{f}$ form an orthogonal set of basis vectors, and projecting equations 6.11 and 6.12 onto these bases yields

$$\vec{f} \bullet \vec{A}^{[H]} = 0 \tag{6.15}$$

[3] For the case $(f_2)^2 + (f_3)^2 = 0$, $\hat{s}$ could be just as well defined as any unit vector orthogonal to $\hat{e}_1$. Definition 6.14 makes the choice independent of the coordinate bases $\hat{e}_2$ and $\hat{e}_3$.

$$\vec{f} \bullet \vec{A}^{[E]} = 0 \tag{6.16}$$

$$(\hat{s} \times \vec{f}) \bullet \vec{A}^{[E]} = \tfrac{1}{\lambda} \hat{s} \bullet \vec{A}^{[H]} \tag{6.17}$$

$$(\hat{s} \times \vec{f}) \bullet \vec{A}^{[H]} = -\tfrac{1}{\lambda} \varepsilon^{[c]} \hat{s} \bullet \vec{A}^{[E]} \tag{6.18}$$

$$(\vec{f} \bullet \vec{f})(\hat{s} \bullet \vec{A}^{[E]}) = -\tfrac{1}{\lambda} (\hat{s} \times \vec{f}) \bullet \vec{A}^{[H]} \tag{6.19}$$

$$(\vec{f} \bullet \vec{f})(\hat{s} \bullet \vec{A}^{[H]}) = \tfrac{1}{\lambda} \varepsilon^{[c]} (\hat{s} \times \vec{f}) \bullet \vec{A}^{[E]} \tag{6.20}$$

By substituting equations 6.17 and 6.18 on the right sides of 6.20 and 6.19, the following condition is obtained,

$$\vec{f} \bullet \vec{f} = \varepsilon^{[c]} / \lambda^2 \tag{6.21}$$

This condition determines the possible values for f_1,

$$f_1 = \pm \sqrt{\frac{\varepsilon^{[c]}}{\lambda^2} - (f_2)^2 - (f_3)^2} \tag{6.22}$$

(It is assumed here that f_1 is nonzero, although this limitation will later be removed.)

Defining the unit vector

$$\hat{p} = \pm \hat{s} \times \left(\frac{\lambda}{\sqrt{\varepsilon^{[c]}}} \vec{f} \right) \tag{6.23}$$

$\vec{A}^{[E]}$ and $\vec{A}^{[H]}$ can be represented as

$$\vec{A}^{[E]} = \hat{s}\, A_s^{[E]} + \hat{p}\, A_p^{[E]} \tag{6.24}$$

$$\vec{A}^{[H]} = \pm(\hat{s}\, A_s^{[H]} + \hat{p}\, A_p^{[H]}) \tag{6.25}$$

This representation automatically satisfies equations 6.15 and 6.16, and equations 6.17 and 6.18 reduce to

$$A_s^{[H]} = \sqrt{\varepsilon^{[c]}}\, A_p^{[E]} \tag{6.26}$$

$$A_p^{[H]} = -\sqrt{\varepsilon^{[c]}}\, A_s^{[E]} \tag{6.27}$$

The "$\pm$" signs in equations 6.23 and 6.25 are, by definition, correlated with the f_1 sign in equation 6.22. The definitions include these signs as a notational convenience, to make the equations symmetric with respect to inversion of the x_1 coordinate ($x_1 \leftarrow -x_1$). The x_1 inversion changes the sign of all vectors' $\hat{e}_1$ projections (e.g., f_1 changes sign, cf. equation 6.8, and consequently expressions 6.23 and 6.25 also change sign). Normally, the equations involving cross products would only be valid in a right-handed coordinate system, whereas the coordinate inversion makes the coordinate system left-handed. However, the equations can be preserved by changing the signs of all cross products in the left-handed system. The $\pm$ sign in equation 6.23 makes this equation valid under x_1 inversion, and the $\pm$ sign in equation 6.25 implicitly changes the sign of $\vec{A}^{[H]}$ under x_1 inversion (without changing $A_s^{[H]}$ or $A_p^{[H]}$) so that equations 6.11 and 6.12 remain valid.

A plane wave can be fully specified by its $\vec{E}$ field because its $\vec{H}$ field is determined by equations 6.26 and 6.27. However, an alternative approach that will be more convenient in the analyses to follow is to specify the electromagnetic field in terms of the $\vec{E}$ and $\vec{H}$ fields' $\hat{s}$ projections ($A_s^{[E]}$ and $A_s^{[H]}$), and let the $\hat{p}$ projections ($A_p^{[E]}$ and $A_p^{[H]}$) be determined implicitly from equations 6.26 and 6.27. This approach is advantageous because continuity conditions between strata are applied to the surface-tangential projections of $\vec{E}$ and $\vec{H}$, of which the fields' $\hat{s}$ projections are a component. The surface-tangential projections also include the projections onto vector $\hat{q}$, which are defined as

$$\hat{q} = \pm \hat{s} \times \hat{e}_1 = \pm(\hat{e}_2\, s_3 - \hat{e}_3\, s_2) \tag{6.28}$$

($\hat{q}$ is a surface-tangential unit vector orthogonal to $\hat{s}$, and a "$\pm$" sign is again included to make the definition invariant under x_1 sign inversion.) Using the following relation,

$$\hat{q} \bullet \hat{p} = \frac{\lambda}{\sqrt{\varepsilon^{[c]}}} f_1 \tag{6.29}$$

(from equations 6.23 and 6.28) the following expressions for the surface-tangential fields are obtained from equations 6.24-27,

$$\hat{s} \bullet \vec{A}^{[E]} = A_s^{[E]} \tag{6.30}$$

$$\hat{s} \bullet \vec{A}^{[H]} = \pm A_s^{[H]} \tag{6.31}$$

$$\hat{q} \bullet \vec{A}^{[E]} = \frac{\lambda}{\varepsilon^{[c]}} f_1 \, A_s^{[H]} \tag{6.32}$$

$$\hat{q} \bullet \vec{A}^{[H]} = \mp \lambda \, f_1 \, A_s^{[E]} \tag{6.33}$$

The above plane-wave equations can be rephrased in the context of general solutions of equations 6.4 and 6.5. $\vec{f}$ is rewritten as $\vec{f}^{[\pm]}[m_1, m_2]$ to indicate explicitly its dependence on the diffraction order indices and the sign choice in equation 6.22,

$$\vec{f}^{[\pm]}[m_1, m_2] = \hat{e}_1 \, f_1^{[\pm]}[m_1, m_2] + \hat{e}_2 \, f_2^{[\pm]}[m_1, m_2] + \hat{e}_3 \, f_3^{(\pm)}[m_1, m_2] \tag{6.34}$$

$$f_1^{[\pm]}[m_1, m_2] = \pm \sqrt{\frac{\varepsilon^{[c]}}{\lambda^2} - (f_2^{[p]}[m_1, m_2])^2 - (f_3^{[p]}[m_1, m_2])^2} \tag{6.35}$$

$$f_2^{[+]}[m_1, m_2] = f_2^{[-]}[m_1, m_2] = f_2^{[p]}[m_1, m_2] \tag{6.36}$$

$$f_3^{[+]}[m_1, m_2] = f_3^{[-]}[m_1, m_2] = f_3^{[p]}[m_1, m_2] \tag{6.37}$$

(cf. equations 6.9, 6.10, and 6.22). $\hat{s}$, $\hat{p}$ and $\hat{q}$ are similarly rewritten as $\hat{s}[m_1, m_2]$, $\hat{p}^{[\pm]}[m_1, m_2]$ and $\hat{q}^{[\pm]}[m_1, m_2]$,

$$\hat{s}[m_1, m_2] = \hat{e}_2 \, s_2[m_1, m_2] + \hat{e}_3 \, s_3[m_1, m_2] \tag{6.38}$$

$$\left(s_2[m_1, m_2], \quad s_3[m_1, m_2]\right) = \\ \begin{cases} \dfrac{\left(-f_3^{[p]}[m_1, m_2], \quad f_2^{[p]}[m_1, m_2]\right)}{\sqrt{(f_2^{[p]}[m_1, m_2])^2 + (f_3^{[p]}[m_1, m_2])^2}} & \text{if } (f_2^{[p]}[m_1, m_2])^2 + (f_3^{[p]}[m_1, m_2])^2 \neq 0 \\ \dfrac{\left(-f_{3,1}^{[g]}, \quad f_{2,1}^{[g]}\right)}{\sqrt{(f_{2,1}^{[g]})^2 + (f_{3,1}^{[g]})^2}} & \text{if } (f_2^{[p]}[m_1, m_2])^2 + (f_3^{[p]}[m_1, m_2])^2 = 0 \end{cases} \tag{6.39}$$

$$\hat{p}^{[\pm]}[m_1, m_2] = \pm \hat{s}[m_1, m_2] \times \left(\frac{\lambda}{\sqrt{\varepsilon^{[c]}}} \vec{f}^{[\pm]}[m_1, m_2] \right) \tag{6.40}$$

$$\hat{q}^{[\pm]}[m_1, m_2] = \pm \hat{s}[m_1, m_2] \times \hat{e}_1 = \pm (\hat{e}_2 \, s_3[m_1, m_2] - \hat{e}_3 \, s_2[m_1, m_2]) \tag{6.41}$$

(cf. equations 6.13, 6.14, 6.23, and 6.28). General solutions of equations 6.4 and 6.5 comprise “up” and “down” waves, which are identified by “+” and “–” labels,

$$ff\vec{E}[m_1, m_2][x_1] = ff\vec{E}^{[+]}[m_1, m_2][x_1] + ff\vec{E}^{[-]}[m_1, m_2][x_1] \tag{6.42}$$

$$ff\vec{H}[m_1,m_2][x_1] = ff\vec{H}^{[+]}[m_1,m_2][x_1] + ff\vec{H}^{[-]}[m_1,m_2][x_1] \tag{6.43}$$

wherein $ff\vec{E}^{[\pm]}[m_1,m_2]$ and $ff\vec{H}^{[\pm]}[m_1,m_2]$ have the functional form

$$ff\vec{E}^{[\pm]}[m_1,m_2][x_1] = \vec{A}^{[E,\pm]}[m_1,m_2]\exp[i\,2\pi\, f_1^{[\pm]}[m_1,m_2]\,x_1] \tag{6.44}$$

$$ff\vec{H}^{[\pm]}[m_1,m_2][x_1] = \vec{A}^{[H,\pm]}[m_1,m_2]\exp[i\,2\pi\, f_1^{[\pm]}[m_1,m_2]\,x_1] \tag{6.45}$$

(cf. equations 6.6 and 6.7). These functions can be represented in terms of their $\hat{s}$ and $\hat{p}$ projections,

$$\begin{aligned} &ff\vec{E}^{[\pm]}[m_1,m_2] = \\ &\quad \hat{s}[m_1,m_2]\, ffE_s^{[\pm]}[m_1,m_2] + \hat{p}^{[\pm]}[m_1,m_2]\, ffE_p^{[\pm]}[m_1,m_2] \end{aligned} \tag{6.46}$$

$$\begin{aligned} &ff\vec{H}^{[\pm]}[m_1,m_2] = \\ &\quad \pm(\hat{s}[m_1,m_2]\, ffH_s^{[\pm]}[m_1,m_2] + \hat{p}^{[\pm]}[m_1,m_2]\, ffH_p^{[\pm]}[m_1,m_2]) \end{aligned} \tag{6.47}$$

(As with equation 6.25, the $\pm$ sign in equation 6.47 is included as a notational convenience, to preserve invariance of cross-product relationships under x_1 sign inversion.) The $\hat{s}$ and $\hat{p}$ projections satisfy the following relationships,

$$ffH_s^{[\pm]}[m_1,m_2] = \sqrt{\varepsilon^{[c]}}\; ffE_p^{[\pm]}[m_1,m_2] \tag{6.48}$$

$$ffH_p^{[\pm]}[m_1,m_2] = -\sqrt{\varepsilon^{[c]}}\; ffE_s^{[\pm]}[m_1,m_2] \tag{6.49}$$

(cf. equations 6.26 and 6.27); hence only the $\hat{s}$ projections need be specified. Surface continuity conditions between strata apply to the fields' projections onto $\hat{s}$ and $\hat{q}$, which are defined as

$$\begin{aligned} &\hat{s}[m_1,m_2]\bullet ff\vec{E}[m_1,m_2][x_1] = \\ &\quad ffE_s^{[+]}[m_1,m_2][x_1] + ffE_s^{[-]}[m_1,m_2][x_1] \end{aligned} \tag{6.50}$$

$$\begin{aligned} &\hat{s}[m_1,m_2]\bullet ff\vec{H}[m_1,m_2][x_1] = \\ &\quad ffH_s^{[+]}[m_1,m_2][x_1] - ffH_s^{[-]}[m_1,m_2][x_1] \end{aligned} \tag{6.51}$$

$$\hat{q}^{[+]}[m_1,m_2]\bullet ff\vec{E}[m_1,m_2][x_1]= \\ \frac{\lambda}{\varepsilon^{[c]}} f_1^{[+]}[m_1,m_2](ffH_s^{[+]}[m_1,m_2][x_1]+ffH_s^{[-]}[m_1,m_2][x_1]) \tag{6.52}$$

$$\hat{q}^{[+]}[m_1,m_2]\bullet ff\vec{H}[m_1,m_2][x_1]= \\ -\lambda\, f_1^{[+]}[m_1,m_2](ffE_s^{[+]}[m_1,m_2][x_1]-ffE_s^{[-]}[m_1,m_2][x_1]) \tag{6.53}$$

(cf. equations 6.30-33).

7. S matrices

Although the plane-wave equations of section 6 only apply in a homogeneous region, they can be formally adopted to apply throughout the grating structure. We can suppose that an infinitesimally thin, homogeneous layer of permittivity $\varepsilon^{[c]}$ is interposed in the grating structure at any particular x_1 level. The layer is too thin to significantly affect the electromagnetic field outside the layer, but within the layer the field comprises up and down waves, as described above. (The up/down field decomposition may lead to numerical indeterminacy or instability when $f_1^{[\pm]}[m_1,m_2]$ is zero or close to zero for some particular order – see definition 6.35 – but this possibility can be avoided by defining $\varepsilon^{[c]}$ to have a positive imaginary part.)

The electromagnetic field transformation across a plane-bounded region sandwiched between two homogeneous layers is described in terms of an “S matrix”, which represents the linear relationship between the fields entering and exiting the region. (“S” connotes “scattering”.) This is illustrated conceptually in Figure 6. In this figure, $F^{[+]}[x_1]$ represents a column vector that includes all of the up-wave amplitudes $ffE_s^{[+]}[m_1,m_2][x_1]$ and $ffH_s^{[+]}[m_1,m_2][x_1]$, and $F^{(-)}[x_1]$ is a column vector including all of the down-wave amplitudes $ffE_s^{[-]}[m_1,m_2][x_1]$ and $ffH_s^{[-]}[m_1,m_2][x_1]$. (The ordering of the amplitudes in $F^{[+]}$ and $F^{[-]}$ will be described in section 9.) The S matrix characterizes a grating stratum in the x_1 interval $x_1^{[0]} < x_1 < x_1^{[1]}$, and it defines a linear mapping between fields at the stratum boundaries. The fields entering the region include the up waves entering from the bottom ($F^{[+]}[x_1^{[0]}]$) and the down waves entering from the top ($F^{[-]}[x_1^{[1]}]$); and the fields exiting the region include the down waves at the bottom ($F^{[-]}[x_1^{[0]}]$) and the up waves at the top ($F^{[+]}[x_1^{[1]}]$). The S matrix comprises submatrices $S00$, $S01$, $S10$, and $S11$, and the mapping between the entering and exiting fields is

$$\begin{pmatrix} F^{[-]}[x_1^{[0]}] \\ F^{[+]}[x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} S00 & S01 \\ S10 & S11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[1]}] \end{pmatrix} \tag{7.1}$$

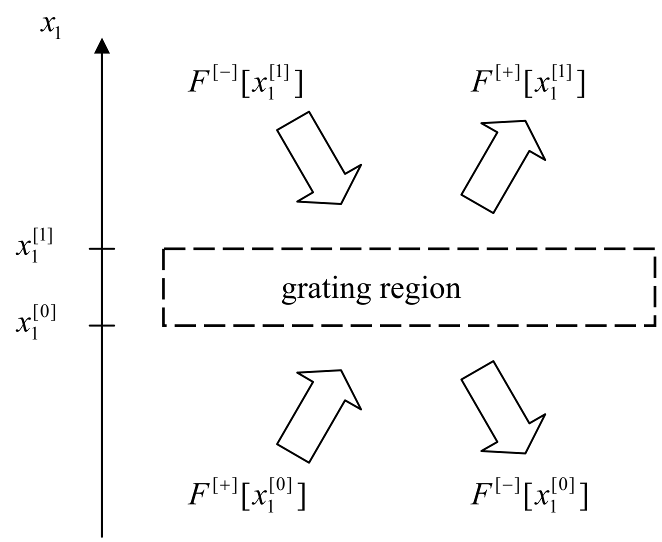

Figure 6. Up and down waves entering and exiting a grating region.

The grating analysis represents all stratum interfaces as infinitesimally thin, homogeneous layers ("fictitious layers"), with the permittivity constant ($\varepsilon^{[c]}$ in section 6) denoted as $\varepsilon^{[f]}$ in the fictitious layers. (The "f" superscript connotes "fictitious".) Separate S matrices are computed for all strata and for the substrate and superstrate boundary interfaces. A "stacking" algorithm described in section 8 is used to combine these S matrices into a composite S matrix for the entire grating structure (i.e., with $x_1^{[0]} = b_1[0] - 0$ and $x_1^{[1]} = b_1[L_1] + 0$ in equation 7.1, wherein "-0" and "$+0$" mean x_1 limits from below and above, respectively, and the b_1 values define the grating boundaries as defined in equations 3.2 and 3.3).

The stacking is performed from bottom to top. The initial value of the S matrix below the grating (i.e., with $x_1^{[1]} = x_1^{[0]} = b_1[0] - 0$) is given by the condition

$$x_1^{[1]} = x_1^{[0]} \quad \rightarrow \quad S00 = S11 = \boldsymbol{0}, \quad S01 = S10 = \boldsymbol{I} \tag{7.2}$$

wherein "$\boldsymbol{0}$" is a zero matrix and "$\boldsymbol{I}$" is an identity matrix. The stacking operations progressively move $x_1^{[1]}$ up through the grating until $x_1^{[1]} = b_1[L_1] + 0$. At each stage of the stacking operation only $S01$ and $S11$ need to be calculated because $F^{[+]}[x_1^{[0]}]$ is zero at $x_1^{[0]} = b_1[0] - 0$ (i.e., there is no incident field entering the grating from the substrate side).

The S matrix computation for an individual stratum is simplified somewhat by taking advantage of the following symmetry relations,

$$\left.\begin{aligned} S00 &= S11 \\ S10 &= S01 \end{aligned}\right\} \text{ for a stratum} \tag{7.3}$$

This is a consequence of the x_1-independence of the permittivity within the stratum, and the identical permittivities $\varepsilon^{[f]}$ above and below the stratum. Also, this condition relies on the magnetic field sign convention represented by the "$\pm$" in equations 6.25 and 6.47. (Relations 7.3 are proved in Appendix C.)

The stacking operation described above can be used to efficiently compute the diffracted field outside the grating, but if the field inside the grating must be determined a slightly different approach, outlined in Appendix D, may be used.

The S-matrix blocks $S00$, $S01$, $S10$, and $S11$ are diagonal or block-diagonal for some stratum types, and section 9 outlines how the S matrix calculations are structured to take advantage of the matrix sparsity. Sections 10 through 16 derive algorithms for computing S matrices specifically for boundary surfaces, coordinate breaks, homogeneous strata, uniperiodic strata, and biperiodic strata.

8. S matrix stacking

The S-matrices across two adjacent x_1 intervals can be "stacked" to determine a composite S matrix covering both intervals, as follows. Consider an S matrix Sa covering the interval from $x_1 = x_1^{[0]}$ to $x_1 = x_1^{[1]}$, and an S matrix Sb covering the interval from $x_1 = x_1^{[1]}$ to $x_1 = x_1^{[2]}$,

$$\begin{pmatrix} F^{[-]}[x_1^{[0]}] \\ F^{[+]}[x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} Sa00 & Sa01 \\ Sa10 & Sa11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[1]}] \end{pmatrix} \tag{8.1}$$

$$\begin{pmatrix} F^{[-]}[x_1^{[1]}] \\ F^{[+]}[x_1^{[2]}] \end{pmatrix} = \begin{pmatrix} Sb00 & Sb01 \\ Sb10 & Sb11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[1]}] \\ F^{[-]}[x_1^{[2]}] \end{pmatrix} \tag{8.2}$$

(cf. equation 7.1). We wish to obtain a composite S matrix covering the combined interval from $x_1 = x_1^{[0]}$ to $x_1 = x_1^{[2]}$,

$$\begin{pmatrix} F^{[-]}[x_1^{[0]}] \\ F^{[+]}[x_1^{[2]}] \end{pmatrix} = \begin{pmatrix} S00 & S01 \\ S10 & S11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[2]}] \end{pmatrix} \tag{8.3}$$

Equations 8.1 and 8.2 can be stated equivalently as

$$\begin{pmatrix} \boldsymbol{0} & Sa01 \\ \boldsymbol{I} & -Sa11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[1]}] \\ F^{[-]}[x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} -Sa00 & \boldsymbol{I} \\ Sa10 & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[0]}] \end{pmatrix} \tag{8.4}$$

$$\begin{pmatrix} \boldsymbol{0} & Sb01 \\ \boldsymbol{I} & -Sb11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[2]}] \\ F^{[-]}[x_1^{[2]}] \end{pmatrix} = \begin{pmatrix} -Sb00 & \boldsymbol{I} \\ Sb10 & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[1]}] \\ F^{[-]}[x_1^{[1]}] \end{pmatrix} \tag{8.5}$$

$F^{[\pm]}[x_1^{[1]}]$ can be eliminated from these equations,

$$\begin{aligned} &\begin{pmatrix} Sa11\, Sa01^{-1} & \boldsymbol{I} \\ Sa01^{-1} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} -Sa00 & \boldsymbol{I} \\ Sa10 & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[0]}] \end{pmatrix} = \begin{pmatrix} F^{(+)}[x_1^{[1]}] \\ F^{(-)}[x_1^{[1]}] \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{0} & Sb10^{-1} \\ \boldsymbol{I} & Sb00\, Sb10^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{0} & Sb01 \\ \boldsymbol{I} & -Sb11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[2]}] \\ F^{[-]}[x_1^{[2]}] \end{pmatrix} \end{aligned} \tag{8.6}$$

Next, the incoming and outgoing field amplitudes are separated on opposite sides of the equation,

$$\begin{aligned} &\begin{pmatrix} -Sa11 & \boldsymbol{I} \\ \boldsymbol{I} & -Sb00 \end{pmatrix} \begin{pmatrix} Sa01^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & Sb10^{-1} \end{pmatrix} \begin{pmatrix} F^{[-]}[x_1^{[0]}] \\ F^{[+]}[x_1^{[2]}] \end{pmatrix} = \\ &\left(\begin{pmatrix} Sa10 & \boldsymbol{0} \\ \boldsymbol{0} & Sb01 \end{pmatrix} + \begin{pmatrix} -Sa11 & \boldsymbol{I} \\ \boldsymbol{I} & -Sb00 \end{pmatrix} \begin{pmatrix} Sa01^{-1}\, Sa00 & \boldsymbol{0} \\ \boldsymbol{0} & Sb10^{-1}\, Sb11 \end{pmatrix} \right) \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[2]}] \end{pmatrix} \end{aligned} \tag{8.7}$$

Comparing this with equation 8.3, the combined S matrix is[4]

$$\begin{aligned} \begin{pmatrix} S00 & S01 \\ S10 & S11 \end{pmatrix} &= \left(\begin{pmatrix} Sa01 & \boldsymbol{0} \\ \boldsymbol{0} & Sb10 \end{pmatrix} \begin{pmatrix} -Sa11 & \boldsymbol{I} \\ \boldsymbol{I} & -Sb00 \end{pmatrix}^{-1} \right. \\ &\qquad \left. \left(\begin{pmatrix} Sa10 & \boldsymbol{0} \\ \boldsymbol{0} & Sb01 \end{pmatrix} + \begin{pmatrix} -Sa11 & \boldsymbol{I} \\ \boldsymbol{I} & -Sb00 \end{pmatrix} \begin{pmatrix} Sa01^{-1}\, Sa00 & \boldsymbol{0} \\ \boldsymbol{0} & Sb10^{-1}\, Sb11 \end{pmatrix} \right) \right) \\ &= \begin{pmatrix} Sa00 + Sa01\, Sb00\, (\boldsymbol{I} - Sa11\, Sb00)^{-1}\, Sa10 & Sa01 (\boldsymbol{I} - Sb00\, Sa11)^{-1}\, Sb01 \\ Sb10\, (\boldsymbol{I} - Sa11\, Sb00)^{-1}\, Sa10 & Sb11 + Sb10\, Sa11 (\boldsymbol{I} - Sb00\, Sa11)^{-1}\, Sb01 \end{pmatrix} \end{aligned} \tag{8.8}$$

[4] Slightly different but equivalent forms of equation 8.8 can be obtained by using the identities

$$Sb00\, (\boldsymbol{I} - Sa11\, Sb00)^{-1} = (\boldsymbol{I} - Sb00\, Sa11)^{-1}\, Sb00$$

$$Sa11 (\boldsymbol{I} - Sb00\, Sa11)^{-1} = (\boldsymbol{I} - Sa11\, Sb00)^{-1}\, Sa11$$

As noted in section 7, only $S01$ and $S11$ need to be calculated when the S matrix stacking is performed from bottom to top. (Note that in equation 8.8 $S01$ and $S11$ do not depend on $Sa00$ or $Sa10$. Thus, $Sa00$ and $Sa10$ need not be specified, and $S00$ and $S10$ are not calculated.) If $S00$ and $S10$ are required, they can be efficiently calculated using top-to-bottom stacking operations (because $S00$ and $S10$ do not depend on $Sb01$ or $Sb11$).

9. Order enumeration and index block partitioning

In principle, the $F^{[\pm]}$ column vectors in equation 7.1 contain an infinite number of elements including all of the wave amplitudes $f\!fE_s^{[\pm]}[m_1, m_2]$ and $f\!fH_s^{[\pm]}[m_1, m_2]$ for diffraction order indices m_1 and m_2 ranging from $-\infty$ to ∞. Each element of $F^{[\pm]}$ corresponds to an index triplet (m_1, m_2, P), which includes the order indices and a "polarization index" P, which can be taken to be 0 for E and 1 for H. In practice, the order indices are limited to a finite set, as defined by conditions 4.11 and 4.12. Furthermore, the set of index triplets can often be partitioned into a number of subsets, which will be termed "decoupled index blocks", such that the exiting wave amplitudes corresponding to each subset depend only on entering wave amplitudes of the same subset. In other words, the S-matrix blocks $S00$, $S01$, $S10$, and $S11$ comprise decoupled submatrices (i.e., $S00$, $S01$, $S10$, and $S11$ are block-diagonal or can be made block-diagonal by applying a permutation to the row and column indices), and each index block corresponds to one of the decoupled submatrices' row and column indices.

The set of all index triplets (m_1, m_2, P) included in the calculation will be organized in terms of two enumeration indices: a block enumeration index j, which labels the decoupled index blocks, and an array index k, which enumerates the (m_1, m_2, P) triplets within each block. This organization is defined by enumeration functions $m_1^{[\text{enum}]}$, $m_2^{[\text{enum}]}$, and $P^{[\text{enum}]}$,

$$
\begin{aligned}
&(m_1, m_2, P) = (m_1^{[\text{enum}]}[j,k], m_2^{[\text{enum}]}[j,k], P^{[\text{enum}]}[j,k]),\\
&\quad j = 1\ldots j^{[\max]}, \quad k = 1\ldots k^{[\max]}[j]
\end{aligned}
\tag{9.1}
$$

The number of decoupled index blocks is $j^{[\max]}$, and the number of (m_1, m_2, P) triplets within block j is $k^{[\max]}[j]$. The component of $F^{[\pm]}[x_1]$ corresponding to block enumeration index j and array index k will be indicated as $F_k^{[\pm, j]}[x_1]$, which is defined as

$$
F_k^{[\pm, j]}[x_1] = \begin{cases} f\!fE_s^{[\pm]}[m_1^{[\text{enum}]}[j,k], m_2^{[\text{enum}]}[j,k]][x_1] & \text{if } P^{[\text{enum}]}[j,k] = 0 \\ f\!fH_s^{[\pm]}[m_1^{[\text{enum}]}[j,k], m_2^{[\text{enum}]}[j,k]][x_1] & \text{if } P^{[\text{enum}]}[j,k] = 1 \end{cases}
\tag{9.2}
$$

The decoupled submatrices in S will be similarly labeled with a " j " superscript, and equation 7.1 will be restated as

$$\begin{pmatrix} F^{[-,j]}[x_1^{[0]}] \\ F^{[+,j]}[x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} S00^{[j]} & S01^{[j]} \\ S10^{[j]} & S11^{[j]} \end{pmatrix} \begin{pmatrix} F^{[+,j]}[x_1^{[0]}] \\ F^{[-,j]}[x_1^{[1]}] \end{pmatrix} \tag{9.3}$$

(for decoupled index block j), or more explicitly,

$$\begin{pmatrix} F_k^{[-,j]}[x_1^{[0]}] \\ F_k^{[+,j]}[x_1^{[1]}] \end{pmatrix} = \sum_{k'=1}^{k^{[\max]}[j]} \begin{pmatrix} S00_{k,k'}^{[j]} & S01_{k,k'}^{[j]} \\ S10_{k,k'}^{[j]} & S11_{k,k'}^{[j]} \end{pmatrix} \begin{pmatrix} F_{k'}^{[+,j]}[x_1^{[0]}] \\ F_{k'}^{[-,j]}[x_1^{[1]}] \end{pmatrix}; \quad k = 1 \ldots k^{[\max]}[j] \tag{9.4}$$

In applying the stacking operation (equation 8.8) to combine S matrices Sa and Sb into a single matrix S, the decoupled index blocks of Sa and Sb are merged in such a way that if a block of Sa intersects a block of Sb, then the union of the two blocks is included in a single block of S.

10. The S matrix for a boundary surface

In defining the S matrix for a boundary surface between two homogeneous regions, $x_1^{[0]}$ is immediately below the surface and $x_1^{[1]}$ is immediately above it in equation 7.1; see Figure 7. (The difference $x_1^{[1]} - x_1^{[0]}$ is infinitesimal.) The permittivity below the surface will be denoted $\varepsilon^{[0]}$, and the permittivity above the surface will be denoted $\varepsilon^{[1]}$.

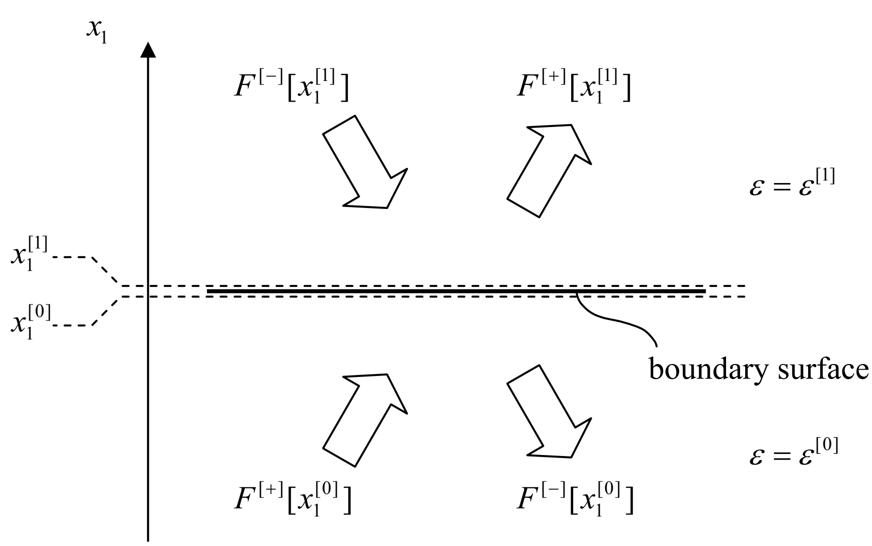

Figure 7. Up and down waves crossing a boundary surface.

The $\vec{E}$ and $\vec{H}$ fields' surface-tangential projections are continuous across the boundary surface. These projections are defined by equations 6.50-53 for Fourier order (m_1, m_2). In equations 6.52 and 6.53 $\varepsilon^{[c]}$ is equal to $\varepsilon^{[0]}$ below the surface and is $\varepsilon^{[1]}$ above the surface. Also, the definition of $f_1^{[+]}$ (equation 6.35) depends on $\varepsilon^{[c]}$, so the two corresponding $f_1^{[+]}$ values will be indicated as $f_1^{[+,0]}$ and $f_1^{[+,1]}$.

$$f_1^{[+,0]}[m_1,m_2] = +\sqrt{\frac{\varepsilon^{[0]}}{\lambda^2} - (f_2^{[p]}[m_1,m_2])^2 - (f_3^{[p]}[m_1,m_2])^2} \tag{10.1}$$

$$f_1^{[+,1]}[m_1,m_2] = +\sqrt{\frac{\varepsilon^{[1]}}{\lambda^2} - (f_2^{[p]}[m_1,m_2])^2 - (f_3^{[p]}[m_1,m_2])^2} \tag{10.2}$$

(In the present context, either of the above values may be zero.) With these notational substitutions, the continuity conditions are

$$\begin{aligned} &ffE_s^{[+]}[m_1,m_2][x_1^{[1]}] + ffE_s^{[-]}[m_1,m_2][x_1^{[1]}] = \\ &\quad ffE_s^{[+]}[m_1,m_2][x_1^{[0]}] + ffE_s^{[-]}[m_1,m_2][x_1^{[0]}] \end{aligned} \tag{10.3}$$

$$\begin{aligned} &ffH_s^{[+]}[m_1,m_2][x_1^{[1]}] - ffH_s^{[-]}[m_1,m_2][x_1^{[1]}] = \\ &\quad ffH_s^{[+]}[m_1,m_2][x_1^{[0]}] - ffH_s^{[-]}[m_1,m_2][x_1^{[0]}] \end{aligned} \tag{10.4}$$

$$\begin{aligned} &\frac{f_1^{[+,1]}[m_1,m_2]}{\varepsilon^{[1]}}(ffH_s^{[+]}[m_1,m_2][x_1^{[1]}] + ffH_s^{[-]}[m_1,m_2][x_1^{[1]}]) = \\ &\quad \frac{f_1^{[+,0]}[m_1,m_2]}{\varepsilon^{[0]}}(ffH_s^{[+]}[m_1,m_2][x_1^{[0]}] + ffH_s^{[-]}[m_1,m_2][x_1^{[0]}]) \end{aligned} \tag{10.5}$$

$$\begin{aligned} &f_1^{[+,1]}[m_1,m_2](ffE_s^{[+]}[m_1,m_2][x_1^{[1]}] - ffE_s^{[-]}[m_1,m_2][x_1^{[1]}]) = \\ &\quad f_1^{[+,0]}[m_1,m_2](ffE_s^{[+]}[m_1,m_2][x_1^{[0]}] - ffE_s^{[-]}[m_1,m_2][x_1^{[0]}]) \end{aligned} \tag{10.6}$$

These conditions can be equivalently restated as following,

$$\begin{pmatrix} ffE_s^{[-]}[m_1,m_2][x_1^{[0]}] \\ ffE_s^{[+]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} = \frac{1}{f_1^{[+,1]}[m_1,m_2] + f_1^{[+,0]}[m_1,m_2]} \cdot \\ \begin{pmatrix} f_1^{[+,0]}[m_1,m_2] - f_1^{[+,1]}[m_1,m_2] & 2\, f_1^{[+,1]}[m_1,m_2] \\ 2\, f_1^{[+,0]}[m_1,m_2] & f_1^{[+,1]}[m_1,m_2] - f_1^{[+,0]}[m_1,m_2] \end{pmatrix} \begin{pmatrix} ffE_s^{[+]}[m_1,m_2][x_1^{[0]}] \\ ffE_s^{[-]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} \tag{10.7}$$

$$\begin{pmatrix} ffH_s^{[-]}[m_1,m_2][x_1^{[0]}] \\ ffH_s^{[+]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} = \frac{1}{\dfrac{f_1^{[+,1]}[m_1,m_2]}{\varepsilon^{[1]}} + \dfrac{f_1^{[+,0]}[m_1,m_2]}{\varepsilon^{[0]}}} \cdot$$
$$\begin{pmatrix} \dfrac{f_1^{[+,1]}[m_1,m_2]}{\varepsilon^{[1]}} - \dfrac{f_1^{[+,0]}[m_1,m_2]}{\varepsilon^{[0]}} & \dfrac{2\, f_1^{[+,1]}[m_1,m_2]}{\varepsilon^{[1]}} \\ \dfrac{2\, f_1^{[+,0]}[m_1,m_2]}{\varepsilon^{[0]}} & \dfrac{f_1^{[+,0]}[m_1,m_2]}{\varepsilon^{[0]}} - \dfrac{f_1^{[+,1]}[m_1,m_2]}{\varepsilon^{[1]}} \end{pmatrix} \begin{pmatrix} ffH_s^{[+]}[m_1,m_2][x_1^{[0]}] \\ ffH_s^{[-]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} \tag{10.8}$$

The boundary surface does not induce any coupling between Fourier orders or between polarization modes; thus the S-matrix blocks $S00$, $S01$, $S10$, and $S11$ are diagonal and $k^{[\max]}[j]$ (the size of the j-th decoupled submatrix) is equal to 1 for all j. With $(m_1,m_2) = (m_1^{[\text{enum}]}[j,1], m_2^{[\text{enum}]}[j,1])$, equation 9.4 takes the form of equation 10.7 for the case $P^{[\text{enum}]}[j,1]=0$, and takes the form of equation 10.8 for the case $P^{[\text{enum}]}[j,1]=1$,

$$S11_{1,1}^{[j]} = \begin{cases} \left(\dfrac{f_1^{[+,1]}[m_1,m_2] - f_1^{[+,0]}[m_1,m_2]}{f_1^{[+,1]}[m_1,m_2] + f_1^{[+,0]}[m_1,m_2]} \right) & \text{if } P^{[\text{enum}]}[j,1]=0 \\ \left(\dfrac{\dfrac{f_1^{[+,0]}[m_1,m_2]}{\varepsilon^{[0]}} - \dfrac{f_1^{[+,1]}[m_1,m_2]}{\varepsilon^{[1]}}}{\dfrac{f_1^{[+,1]}[m_1,m_2]}{\varepsilon^{[1]}} + \dfrac{f_1^{[+,0]}[m_1,m_2]}{\varepsilon^{[0]}}} \right) & \text{if } P^{[\text{enum}]}[j,1]=1 \end{cases} \tag{10.9}$$

$$S01_{1,1}^{[j]} = \begin{cases} 1 + S11_{1,1}^{[j]} & \text{if } P^{[\text{enum}]}[j,1]=0 \\ 1 - S11_{1,1}^{[j]} & \text{if } P^{[\text{enum}]}[j,1]=1 \end{cases} \tag{10.10}$$

$$S00_{1,1}^{[j]} = -S11_{1,1}^{[j]} \tag{10.11}$$

$$S10_{1,1}^{[j]} = \begin{cases} 1 - S11_{1,1}^{[j]} & \text{if } P^{[\text{enum}]}[j,1]=0 \\ 1 + S11_{1,1}^{[j]} & \text{if } P^{[\text{enum}]}[j,1]=1 \end{cases} \tag{10.12}$$

(If the denominator in equation 10.9 is zero, then $S11_{1,1}^{[j]}$ can be set to zero. This condition should not occur unless $\varepsilon^{[0]}$ and $\varepsilon^{[1]}$ are real-valued and equal.)

11. The S matrix for a homogeneous stratum (without surfaces)

The S matrix for a homogeneous stratum (excluding its boundary surfaces) is determined from equations 6.44 and 6.45. These equations imply

$$ffE_s^{[\pm]}[m_1,m_2][x_1^{[1]}] = \exp[\pm i\,\phi]\,ffE_s^{[\pm]}[m_1,m_2][x_1^{[0]}] \tag{11.1}$$

$$ffH_s^{[\pm]}[m_1,m_2][x_1^{[1]}] = \exp[\pm i\,\phi]\,ffH_s^{[\pm]}[m_1,m_2][x_1^{[0]}] \tag{11.2}$$

wherein the phase shift ϕ is

$$\phi = 2\pi\, f_1^{[+]}[m_1,m_2](x_1^{[1]} - x_1^{[0]}) \tag{11.3}$$

The stratum boundaries are at $x_1 = x_1^{[0]}$ and $x_1 = x_1^{[1]} \ge x_1^{[0]}$ (see Figure 8). $f_1^{[+]}[m_1,m_2]$ is defined by equation 6.35. (The stratum permittivity is denoted as $\varepsilon^{[c]}$ in equation 6.35.)

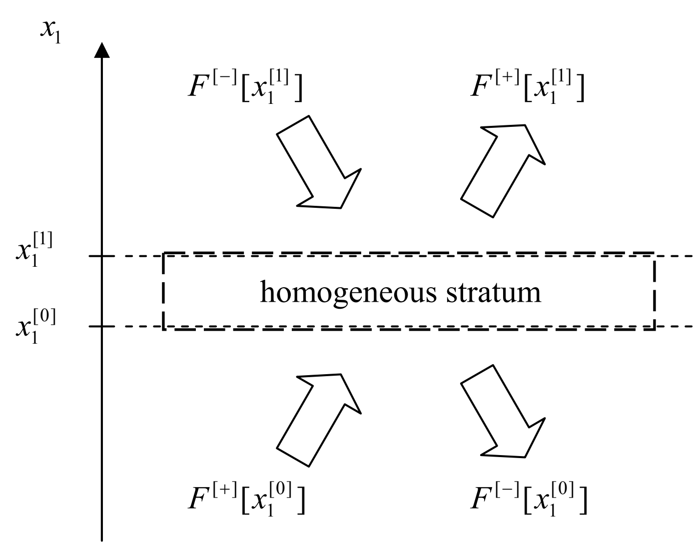

Figure 8. Up and down waves traversing a homogeneous stratum.

Equations 11.1 and 11.2 can be equivalently stated as follows,

$$\begin{pmatrix} ffE_s^{[-]}[m_1,m_2][x_1^{[0]}] \\ ffE_s^{[+]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} 0 & \exp[i\,\phi] \\ \exp[i\,\phi] & 0 \end{pmatrix} \begin{pmatrix} ffE_s^{[+]}[m_1,m_2][x_1^{[0]}] \\ ffE_s^{[-]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} \tag{11.4}$$

$$\begin{pmatrix} ffH_s^{[-]}[m_1,m_2][x_1^{[0]}] \\ ffH_s^{[+]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} 0 & \exp[i\,\phi] \\ \exp[i\,\phi] & 0 \end{pmatrix} \begin{pmatrix} ffH_s^{[+]}[m_1,m_2][x_1^{[0]}] \\ ffH_s^{[-]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} \tag{11.5}$$

There is no diffractive coupling or polarization coupling within the stratum, so $k^{[\max]}[j]=1$ in equation 9.4; and with $(m_1,m_2)=(m_1^{[\mathrm{enum}]}[j,1],m_2^{[\mathrm{enum}]}[j,1])$ and $P^{[\mathrm{enum}]}[j,1]$ = 0 or 1 the S matrix's j-th diagonal elements are

$$S00_{1,1}^{[j]}=S11_{1,1}^{[j]}=0 \tag{11.6}$$

$$S10_{1,1}^{[j]}=S01_{1,1}^{[j]}=\exp[i\,\phi] \tag{11.7}$$

The stratum's S matrix defined by equations 11.6 and 11.7 can be combined with the stratum's boundary-surface S matrices (equations 10.9-12) via stacking (equation 8.8). The resulting composite S matrix can contain indeterminacies when a Fourier order's grating-normal frequency (f_1) is zero in the stratum or in either of its bounding layers. However, this problem can be avoided by describing the layer in terms of an S matrix representing the electromagnetic field transitions between two fictitious layers of permittivity $\varepsilon^{[f]}$ bounding the stratum. (The S matrix includes the bounding surfaces.) S matrices for periodic strata are similarly defined in relation to fictitious bounding layers, and section 16 derives the S matrix for a homogeneous layer as a specialization of the periodic case.

12. The S matrix for a coordinate break

A coordinate break applies a translational offset to the x_2, x_3 coordinates at a particular height x_1 in the grating stack. The coordinate break is represented by an S matrix, as in equation 7.1, wherein $x_1^{[0]}$ represents the x_1 height immediately below the break and $x_1^{[1]}$ represents the height immediately above it. (The difference $x_1^{[1]}-x_1^{[0]}$ is infinitesimal.) The column vectors $F^{[\pm]}[x_1]$ in equation 7.1 represent field quantities evaluated at $(x_2,x_3)=(0,0)$ below the break ($x_1<x_1^{[0]}$), and at $(x_2',x_3')=(0,0)$ above the break ($x_1>x_1^{[1]}$), wherein the primed coordinates are defined by

$$(x_2',x_3')=(x_2,x_3)-(\Delta x_2,\Delta x_3) \tag{12.1}$$

Δx_2 and Δx_3 are the translational offsets. For example, Figure 9 conceptually illustrates a break applied to the x_2 coordinate, with offset Δx_2.

The offset vector $\Delta\vec{x}$ is defined as in equation 3.39,

$$\Delta\vec{x}=\hat{e}_2\,\Delta x_2+\hat{e}_3\,\Delta x_3 \tag{12.2}$$

The electric field representation $\vec{E}[\hat{e}_1\, x_1^{[0]} + \hat{e}_2\, x_2 + \hat{e}_3\, x_3]$ below the break is transformed into the representation $\vec{E}[\hat{e}_1\, x_1^{[1]} + \hat{e}_2\, x_2' + \hat{e}_3\, x_3']$ above the break, wherein

$$\begin{aligned}&\vec{E}[\hat{e}_1\, x_1^{[1]} + \hat{e}_2\, x_2' + \hat{e}_3\, x_3'] = \vec{E}[\hat{e}_1\, x_1^{[0]} + \hat{e}_2\, x_2 + \hat{e}_3\, x_3]\\ &\quad = \vec{E}[\hat{e}_1\, x_1^{[0]} + \hat{e}_2\, x_2' + \hat{e}_3\, x_3' + \Delta\bar{x}]\end{aligned} \tag{12.3}$$

Substituting the field's Fourier expansion 5.7 in 12.3 and using the condition $\vec{f}^{[p]}[m_1, m_2] \bullet \hat{e}_1 = 0$ (from equation 5.9), the following relationship is obtained,

$$\begin{aligned}&\sum_{m_1, m_2} ff\vec{E}[m_1, m_2][x_1^{[1]}] \exp[i\, 2\pi\, \vec{f}^{[p]}[m_1, m_2] \bullet (\hat{e}_2\, x_2' + \hat{e}_3\, x')] =\\ &\quad \sum_{m_1, m_2} ff\vec{E}[m_1, m_2][x_1^{[0]}] \exp[i\, 2\pi\, \vec{f}^{[p]}[m_1, m_2] \bullet (\hat{e}_2\, x_2' + \hat{e}_3\, x_3' + \Delta\bar{x})]\end{aligned} \tag{12.4}$$

Hence the coordinate translation induces the following phase shift in the field's Fourier coefficients,

$$ff\vec{E}[m_1, m_2][x_1^{[1]}] = ff\vec{E}[m_1, m_2][x_1^{[0]}] \exp[i\, \phi] \tag{12.5}$$

wherein

$$\phi = 2\pi\, \vec{f}^{[p]}[m_1, m_2] \bullet \Delta\bar{x} \tag{12.6}$$

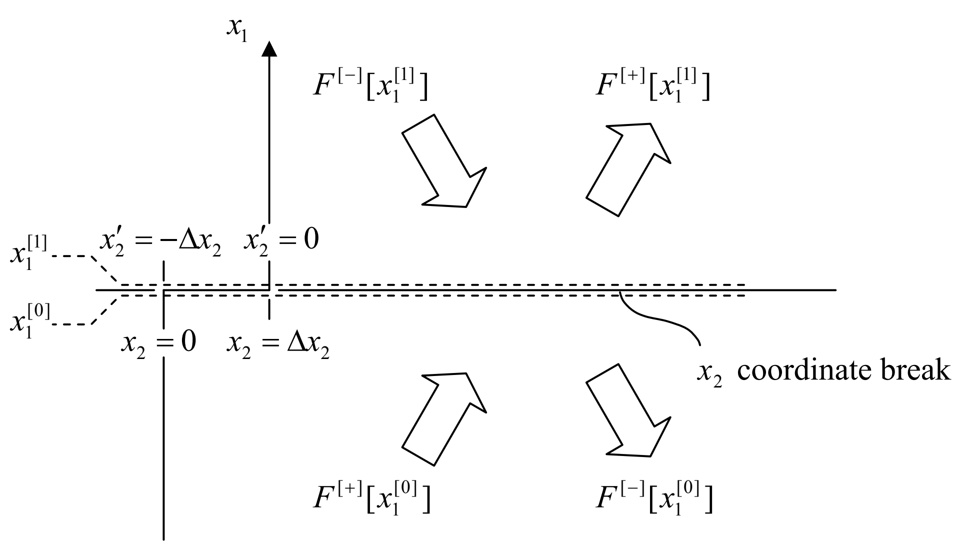

Figure 9. Up and down waves with a coordinate break.

A similar phase shift is applied to the $ff\vec{H}$ coefficients in equation 5.12, and to the $ffE_s^{[\pm]}$ and $ffH_s^{[\pm]}$ functions defined by equations 6.50-53. The phase-shift relations can be expressed as follows,

$$\begin{pmatrix} ffE_s^{[-]}[m_1,m_2][x_1^{[0]}] \\ ffE_s^{[+]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} 0 & \exp[-i\,\phi] \\ \exp[i\,\phi] & 0 \end{pmatrix} \begin{pmatrix} ffE_s^{[+]}[m_1,m_2][x_1^{[0]}] \\ ffE_s^{[-]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} \tag{12.7}$$

$$\begin{pmatrix} ffH_s^{[-]}[m_1,m_2][x_1^{[0]}] \\ ffH_s^{[+]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} 0 & \exp[-i\,\phi] \\ \exp[i\,\phi] & 0 \end{pmatrix} \begin{pmatrix} ffH_s^{[+]}[m_1,m_2][x_1^{[0]}] \\ ffH_s^{[-]}[m_1,m_2][x_1^{[1]}] \end{pmatrix} \tag{12.8}$$

The phase shift does not induce any diffractive coupling or polarization coupling, so $k^{[\max]}[j]=1$ in equation 9.4; and with $(m_1,m_2)=(m_1^{[\text{enum}]}[j,1], m_2^{[\text{enum}]}[j,1])$ and $P^{[\text{enum}]}[j,1]$ = 0 or 1 the S matrix's j-th diagonal elements are

$$S00_{1,1}^{[j]} = S11_{1,1}^{[j]} = 0 \tag{12.9}$$

$$S10_{1,1}^{[j]} = \exp[i\,\phi] \tag{12.10}$$

$$S01_{1,1}^{[j]} = \exp[-i\,\phi] \tag{12.11}$$

13. The S matrix for a biperiodic stratum

As a consequence of the grating periodicity conditions 3.15 and 3.16, the permittivity in stratum l_1 (equation 3.10) can be represented by a Fourier series of the form,

$$\begin{aligned} \varepsilon[\vec{x}] &= \varepsilon 1[l_1][x_2,x_3] \\ &= \sum_{n_1,n_2} ff\varepsilon 1[l_1,n_1,n_2] \exp[i\,2\pi\,(n_1\,\vec{f}_1^{[s]}[l_1] + n_2\,\vec{f}_2^{[s]}[l_1]) \bullet \vec{x}] \end{aligned} \tag{13.1}$$

(The stratum's basis frequencies $\vec{f}_1^{[s]}[l_1]$ and $\vec{f}_2^{[s]}[l_1]$ are defined by equations 3.28-30, and the periodicity conditions 3.15 and 3.16 can be verified directly using equations 13.1 and 3.30.) Appendix A describes the explicit form of the Fourier coefficients $ff\varepsilon 1[l_1,n_1,n_2]$, based on the grating geometry specification outlined in section 3.

The stratum's basis frequencies $\vec{f}_1^{[s]}[l_1]$ and $\vec{f}_2^{[s]}[l_1]$ are linear combinations of the grating's fundamental basis frequencies $\vec{f}_1^{[g]}$ and $\vec{f}_2^{[g]}$,

$$\vec{f}_1^{[s]}[l_1] = h_{1,1}[l_1]\,\vec{f}_1^{[g]} + h_{1,2}[l_1]\,\vec{f}_2^{[g]} \tag{13.2}$$

$$\vec{f}_2^{[s]}[l_1] = h_{2,1}[l_1]\,\vec{f}_1^{[g]} + h_{2,2}[l_1]\,\vec{f}_2^{[g]} \tag{13.3}$$

(cf. equation 3.31). The diffraction order indices (m_1, m_2) of any particular Fourier order in the electromagnetic field expansions 5.7 and 5.12 can be represented as

$$\begin{pmatrix} m_1, & m_2 \end{pmatrix} = \begin{pmatrix} m_1^{[0]}, & m_2^{[0]} \end{pmatrix} + \begin{pmatrix} n_1, & n_2 \end{pmatrix} \begin{pmatrix} h_{1,1}[l_1] & h_{1,2}[l_1] \\ h_{2,1}[l_1] & h_{2,2}[l_1] \end{pmatrix} \tag{13.4}$$

wherein n_1, and n_2 are integers defined by

$$\begin{pmatrix} n_1, & n_2 \end{pmatrix} = \operatorname{floor}\left[\begin{pmatrix} m_1, & m_2 \end{pmatrix} \begin{pmatrix} h_{1,1}[l_1] & h_{1,2}[l_1] \\ h_{2,1}[l_1] & h_{2,2}[l_1] \end{pmatrix}^{-1} \right] \tag{13.5}$$

and $m_1^{[0]}$ and $m_2^{[0]}$ are integers in the range

$$0 \le \frac{m_1^{[0]}\,h_{2,2}[l_1] - m_2^{[0]}\,h_{2,1}[l_1]}{h_{1,1}[l_1]h_{2,2}[l_1] - h_{1,2}[l_1]h_{2,1}[l_1]} < 1 \tag{13.6}$$

$$0 \le \frac{-m_1^{[0]}\,h_{1,2}[l_1] + m_2^{[0]}\,h_{1,1}[l_1]}{h_{1,1}[l_1]h_{2,2}[l_1] - h_{1,2}[l_1]h_{2,1}[l_1]} < 1 \tag{13.7}$$

(The “floor” function rounds toward $-\infty$.) The integer pair $(m_1^{[0]}, m_2^{[0]})$ will be termed the “base indices”, and (n_1, n_2) will be termed “suborder indices”, corresponding to order (m_1, m_2) in stratum l_1.

The set of all order index pairs (m_1, m_2) corresponding to a particular base index pair $(m_1^{[0]}, m_2^{[0]})$ defines a decoupled index block. The set of all base index pairs $(m_1^{[0]}, m_2^{[0]})$ satisfying relations 13.6 and 13.7 will be enumerated and denoted as $(m_1^{[0]}[j], m_2^{[0]}[j])$ for the j-th pair, and the corresponding index block will be designated as the j-th decoupled index block. In equations 5.7 and 5.12 the field amplitudes $\mathit{ff}\vec{E}[m_1, m_2]$ and $\mathit{ff}\vec{H}[m_1, m_2]$ corresponding to index block j and suborder indices (n_1, n_2) will be represented as follows

$$ff\vec{E}^{[j]}[n_1,n_2][x_1] = \\ ff\vec{E}[m_1^{[0]}[j] + n_1\, h_{1,1}[l_1] + n_2\, h_{2,1}[l_1], m_2^{[0]}[j] + n_1\, h_{1,2}[l_1] + n_2\, h_{2,2}[l_1]][x_1] \tag{13.8}$$

$$ff\vec{H}^{[j]}[n_1,n_2][x_1] = \\ ff\vec{H}[m_1^{[0]}[j] + n_1\, h_{1,1}[l_1] + n_2\, h_{2,1}[l_1], m_2^{[0]}[j] + n_1\, h_{1,2}[l_1] + n_2\, h_{2,2}[l_1]][x_1] \tag{13.9}$$

The following grating-tangential spatial frequencies are also defined,

$$\vec{f}^{[ip,j]} = \vec{f}^{[ip]} + m_1^{[0]}[j]\, \vec{f}_1^{[g]} + m_2^{[0]}[j]\, \vec{f}_2^{[g]} \tag{13.10}$$

$$\vec{f}^{[p,j]}[n_1,n_2] = \vec{f}^{[ip,j]} + n_1\, \vec{f}_1^{[s]}[l_1] + n_2\, \vec{f}_2^{[s]}[l_1] \tag{13.11}$$

(cf. equation 5.8). With these definitions, along with definition 5.8, the field expansions 5.7 and 5.12 can be represented as

$$\vec{E}[\vec{x}] = \sum_{j,n_1,n_2} ff\vec{E}^{[j]}[n_1,n_2][x_1] \exp[i\, 2\pi\, \vec{f}^{[p,j]}[n_1,n_2] \bullet \vec{x}] \tag{13.12}$$

$$\vec{H}[\vec{x}] = \sum_{j,n_1,n_2} ff\vec{H}^{[j]}[n_1,n_2][x_1] \exp[i\, 2\pi\, \vec{f}^{[p,j]}[n_1,n_2] \bullet \vec{x}] \tag{13.13}$$

In developing the S matrix for stratum l_1, the $\hat{e}_2$, $\hat{e}_3$ coordinate orientation is chosen so that $\hat{e}_2$ is parallel to $\vec{f}_1^{[s]}[l_1]$,

$$f_{3,1}^{[s]}[l_1] = 0; \quad \vec{f}_1^{[s]}[l_1] = \hat{e}_2\, f_{2,1}^{[s]}[l_1] \tag{13.14}$$

(The S matrix definition is independent of the $\hat{e}_2$, $\hat{e}_3$ coordinate orientation because the definition of $ffE_s^{[\pm]}$ and $ffH_s^{[\pm]}$ implicit in equations 6.46 and 6.47 is coordinate-independent. Thus, the choice of coordinate orientation has no effect on the S matrix.) The permittivity expansion 13.1 then takes the form of a one-dimensional Fourier series in x_3 (the coordinate parallel to the grating stripes),

$$\varepsilon[\vec{x}] = \varepsilon 1[l_1][x_2,x_3] = \sum_{n_2} f\varepsilon 1[l_1,n_2][x_2] \exp[i\, 2\pi\, n_2\, f_{3,2}^{[s]}[l_1]\, x_3] \tag{13.15}$$

wherein each Fourier coefficient $f\varepsilon 1[l_1,n_2][x_2]$ is a one-dimensional Fourier series in x_2 (the stripe-transverse coordinate),

$$f\varepsilon 1[l_1,n_2][x_2]=\sum_{n_1} ff\varepsilon 1[l_1,n_1,n_2]\exp[i\,2\pi\,(n_1\,f_{2,1}^{[s]}[l_1]+n_2\,f_{2,2}^{[s]}[l_1])\,x_2] \tag{13.16}$$

($\varepsilon 1[l_1][x_2,x_3]$ and $f\varepsilon 1[l_1,n_2][x_2]$ are independent of x_2 within each stripe, and $\varepsilon 1[l_1][x_2,x_3]$ is independent of both x_2 and x_3 within each structural block.) The electromagnetic field expansions 13.12 and 13.13 can be similarly represented as nested one-dimensional expansions,

$$\vec{E}[\vec{x}]=\sum_j\sum_{n_2} f\vec{E}^{[j]}[n_2][x_1,x_2]\exp[i\,2\pi\,(f_3^{[ip,j]}+n_2\,f_{3,2}^{[s]}[l_1])\,x_3] \tag{13.17}$$

$$\vec{H}[\vec{x}]=\sum_j\sum_{n_2} f\vec{H}^{[j]}[n_2][x_1,x_2]\exp[i\,2\pi\,(f_3^{[ip,j]}+n_2\,f_{3,2}^{[s]}[l_1])\,x_3] \tag{13.18}$$

wherein

$$f\vec{E}^{[j]}[n_2][x_1,x_2]=\sum_{n_1} ff\vec{E}^{[j]}[n_1,n_2][x_1]\exp[i\,2\pi\,(f_2^{[ip,j]}+n_1\,f_{2,1}^{[s]}[l_1]+n_2\,f_{2,2}^{[s]}[l_1])\,x_2] \tag{13.19}$$

$$f\vec{H}^{[j]}[n_2][x_1,x_2]=\sum_{n_1} ff\vec{H}^{[j]}[n_1,n_2][x_1]\exp[i\,2\pi\,(f_2^{[ip,j]}+n_1\,f_{2,1}^{[s]}[l_1]+n_2\,f_{2,2}^{[s]}[l_1])\,x_2] \tag{13.20}$$

The Fourier expansions in x_3, equations 13.15, 13.17, and 13.18, can be substituted in the Maxwell Equations 6.1 and 6.2, which have the explicit coordinate representation,

$$\partial_2\,E_3-\partial_3\,E_2=i\,\tfrac{2\pi}{\lambda}\,H_1 \tag{13.21}$$

$$\partial_3\,E_1-\partial_1\,E_3=i\,\tfrac{2\pi}{\lambda}\,H_2 \tag{13.22}$$

$$\partial_1\,E_2-\partial_2\,E_1=i\,\tfrac{2\pi}{\lambda}\,H_3 \tag{13.23}$$

$$\partial_2\,H_3-\partial_3\,H_2=-i\,\tfrac{2\pi}{\lambda}\,\varepsilon\,E_1 \tag{13.24}$$

$$\partial_3\,H_1-\partial_1\,H_3=-i\,\tfrac{2\pi}{\lambda}\,\varepsilon\,E_2 \tag{13.25}$$

$$\partial_1 H_2 - \partial_2 H_1 = -i\tfrac{2\pi}{\lambda}\varepsilon E_3 \tag{13.26}$$

wherein ∂_ν represents the derivative with respect to x_ν. In making the substitution, Laurent's product rule for Fourier series should only be applied when the product factors do not have concurrent discontinuities in x_3; otherwise convergence of the Fourier series will be nonuniform and severe numerical instabilities may result [Ref. 3]. The $\vec{E}$ and $\vec{H}$ fields' surface-tangential components are continuous across permittivity discontinuity surfaces, and the grating walls between structural blocks within each stripe are parallel to $\hat{e}_1$ and $\hat{e}_2$; hence the factors E_1 and E_2 in equations 13.24 and 13.25 are continuous with x_3 and there is no problem applying Laurent's rule in these equations. The factors ε and E_3 in equation 13.26 will generally have concurrent discontinuities; however the terms H_1 and H_2 on the left side of the equation are continuous (and hence, so are their tangential derivatives $\partial_2 H_1$ and $\partial_1 H_2$), so Laurent's rule can be reliably applied by moving the ε factor to the left side of the equation,

$$\tfrac{1}{\varepsilon}\left(\partial_1 H_2 - \partial_2 H_1\right) = -i\tfrac{2\pi}{\lambda} E_3 \tag{13.27}$$

The reciprocal permittivity factor in equation 13.27 is represented by a Fourier series similar to equation 13.15,

$$\frac{1}{\varepsilon[\vec{x}]} = \frac{1}{\varepsilon 1[l_1][x_2, x_3]} = \sum_{n_2} fr\varepsilon 1[l_1, n_2][x_2]\exp[i\,2\pi\, n_2\, f_{3,2}^{[\mathrm{s}]}[l_1]\, x_3] \tag{13.28}$$

(The "r" in "$fr\varepsilon 1$" connotes "reciprocal".) Substituting equations 13.17, 13.18, 13.15, and 13.28 in equations 13.21-25 and 13.27 and separating Fourier orders, the following equations are obtained,

$$\partial_2\, fE_3^{[j]}[n_2] - i\,2\pi\,(f_3^{[ip,j]} + n_2\, f_{3,2}^{[\mathrm{s}]}[l_1])\, fE_2^{[j]}[n_2] = i\tfrac{2\pi}{\lambda}\, fH_1^{[j]}[n_2] \tag{13.29}$$

$$i\,2\pi\,(f_3^{[ip,j]} + n_2\, f_{3,2}^{[\mathrm{s}]}[l_1])\, fE_1^{[j]}[n_2] - \partial_1\, fE_3^{[j]}[n_2] = i\tfrac{2\pi}{\lambda}\, fH_2^{[j]}[n_2] \tag{13.30}$$

$$\partial_1\, fE_2^{[j]}[n_2] - \partial_2\, fE_1^{[j]}[n_2] = i\tfrac{2\pi}{\lambda}\, fH_3^{[j]}[n_2] \tag{13.31}$$

$$\begin{aligned} &\partial_2\, fH_3^{[j]}[n_2] - i\,2\pi\,(f_3^{[ip,j]} + n_2\, f_{3,2}^{[\mathrm{s}]}[l_1])\, fH_2^{[j]}[n_2] = \\ &\quad -i\tfrac{2\pi}{\lambda}\sum_{n_2'} f\varepsilon 1[l_1, n_2 - n_2']\, fE_1^{[j]}[n_2'] \end{aligned} \tag{13.32}$$

$$i\,2\pi\,(f_3^{[ip,j]}+n_2\,f_{3,2}^{[s]}[l_1])\,fH_1^{[j]}[n_2]-\partial_1\,fH_3^{[j]}[n_2]=\\ -i\tfrac{2\pi}{\lambda}\sum_{n_2'}f\varepsilon 1[l_1,n_2-n_2']\,fE_2^{[j]}[n_2'] \tag{13.33}$$

$$\sum_{n_2'}fr\varepsilon 1[l_1,n_2-n_2']\left(\partial_1\,fH_2^{[j]}[n_2']-\partial_2\,fH_1^{[j]}[n_2']\right)=\\ -i\tfrac{2\pi}{\lambda}\,fE_3^{[j]}[n_2] \tag{13.34}$$

Next, the Fourier expansions in x_2, equations 13.16, 13.19, and 13.20, can be substituted above. Once again, Laurent's rule is applied to the products inside the summations, taking care to avoid applying the rule to products with concurrent discontinuities. The grating walls between stripes within each stratum are parallel to $\hat{e}_1$ and $\hat{e}_3$; hence the field components E_1, E_3, H_1 and H_3 and their associated Fourier coefficients $fE_1^{[j]}[n_2]$, $fE_3^{[j]}[n_2]$, $fH_1^{[j]}[n_2]$ and $fH_3^{[j]}[n_2]$ (and also the derivative $\partial_1\,fH_3^{[j]}[n_2]$) are all continuous with x_2, whereas terms associated with E_2 and H_2 generally are not. Thus, Laurent's rule can be reliably applied to equation 13.32, but equations 13.33 and 13.34 must be modified to move the permittivity terms to the side of each equation containing the continuous field quantities.

The modification is facilitated by making the following definitions

$$t\varepsilon 1[l_1,n_2,n_2'][x_2]=f\varepsilon 1[l_1,n_2-n_2'][x_2] \tag{13.35}$$

$$tr\varepsilon 1[l_1,n_2,n_2'][x_2]=fr\varepsilon 1[l_1,n_2-n_2'][x_2] \tag{13.36}$$

The "t" prefix connotes a Toeplitz matrix, and corresponding reciprocal matrices $rt\varepsilon 1$ and $rtr\varepsilon 1$ are defined by the conditions

$$\sum_{n_2'}(rt\varepsilon 1[l_1,n_2,n_2'][x_2])\,(t\varepsilon 1[l_1,n_2',n_2''][x_2])=\begin{cases}1, & n_2=n_2''\\ 0, & n_2\neq n_2''\end{cases} \tag{13.37}$$

$$\sum_{n_2'}(rtr\varepsilon 1[l_1,\,n_2,n_2'][x_2])\,(tr\varepsilon 1[l_1,\,n_2',n_2''][x_2])=\begin{cases}1, & n_2=n_2''\\ 0, & n_2\neq n_2''\end{cases} \tag{13.38}$$

(These matrices are independent of x_2 within each stripe.) With these definitions, equations 13.33 and 13.34 can be restated in a form that is suitable for application of Laurent's rule,

$$\sum_{n_2'} rt\varepsilon 1[l_1,n_2,n_2']\left(i\,2\pi\,(f_3^{[ip,j]}+n_2'\,f_{3,2}^{[s]}[l_1])\,fH_1^{[j]}[n_2']-\partial_1\,fH_3^{[j]}[n_2']\right)= \\ -i\,\tfrac{2\pi}{\lambda}\,fE_2^{[j]}[n_2] \tag{13.39}$$

$$\partial_1\,fH_2^{[j]}[n_2]-\partial_2\,fH_1^{[j]}[n_2]= \\ -i\,\tfrac{2\pi}{\lambda}\sum_{n_2'} rtr\varepsilon 1[l_1,n_2,n_2']\,fE_3^{[j]}[n_2'] \tag{13.40}$$

Equation 13.16 defines the Fourier expansion, with respect to x_2, of the permittivity term in equation 13.32. The permittivity terms in equations 13.39 and 13.40 have similar Fourier expansions, which are derived from the periodicity conditions 3.15 and 3.16. It follows from equations 3.30 and 13.14 that equations 3.15 and 3.16 can be equivalently stated

$$\varepsilon 1[l_1][x_2+1/\,f_{2,1}^{[s]}[l_1],x_3-f_{2,2}^{[s]}[l_1]/(f_{2,1}^{[s]}[l_1]\,f_{3,2}^{[s]}[l_1])]=\varepsilon 1[l_1][x_2,x_3] \tag{13.41}$$

$$\varepsilon 1[l_1][x_2,x_3+1/\,f_{3,2}^{[s]}[l_1]]=\varepsilon 1[l_1][x_2,x_3] \tag{13.42}$$

The permittivity's Fourier expansion 13.15 follows from equation 13.42, and based on equations 13.15 and 13.41 the Fourier coefficients satisfy the following x_2-periodicity condition,

$$f\varepsilon 1[l_1,n_2][x_2+1/\,f_{2,1}^{[s]}[l_1]]\exp[-i\,2\pi\,n_2\,f_{2,2}^{[s]}[l_1](x_2+1/\,f_{2,1}^{[s]}[l_1])]= \\ f\varepsilon 1[l_1,n_2][x_2]\exp[-i\,2\pi\,n_2\,f_{2,2}^{[s]}[l_1]\,x_2] \tag{13.43}$$

Expansion 13.16 follows from condition 13.43.

Based on equation 13.43, the Toeplitz matrix $t\varepsilon 1$ defined by equation 13.35 satisfies a similar periodicity condition,

$$t\varepsilon 1[l_1,n_2,n_2'][x_2+1/\,f_{2,1}^{[s]}[l_1]]\exp[-i\,2\pi\,(n_2-n_2')\,f_{2,2}^{[s]}[l_1](x_2+1/\,f_{2,1}^{[s]}[l_1])]= \\ t\varepsilon 1[l_1,n_2,n_2'][x_2]\exp[-i\,2\pi\,(n_2-n_2')\,f_{2,2}^{[s]}[l_1]\,x_2] \tag{13.44}$$

and the reciprocal Toeplitz matrix $rt\varepsilon 1$ defined by condition 13.37 has the same periodicity,

$$rt\varepsilon 1[l_1,n_2,n_2'][x_2+1/\,f_{2,1}^{[s]}[l_1]]\exp[-i\,2\pi\,(n_2-n_2')\,f_{2,2}^{[s]}[l_1](x_2+1/\,f_{2,1}^{[s]}[l_1])]= \\ rt\varepsilon 1[l_1,n_2,n_2'][x_2]\exp[-i\,2\pi\,(n_2-n_2')\,f_{2,2}^{[s]}[l_1]\,x_2] \tag{13.45}$$

Hence, $rt\varepsilon 1$ can be represented by a Fourier expansion similar to equation 13.16,

$$rt\varepsilon 1[l_1,n_2,n_2'][x_2] = \sum_{n_1} frt\varepsilon 1[l_1,n_1,n_2,n_2'] \exp[i\,2\pi\,(n_1\,f_{2,1}^{[s]}[l_1] + (n_2-n_2')\,f_{2,2}^{[s]}[l_1])\,x_2] \tag{13.46}$$

The matrix $rtr\varepsilon 1$ defined by condition 13.38 can be similarly represented as

$$rtr\varepsilon 1[l_1,n_2,n_2'][x_2] = \sum_{n_1} frtr\varepsilon 1[l_1,n_1,n_2,n_2'] \exp[i\,2\pi\,(n_1\,f_{2,1}^{[s]}[l_1] + (n_2-n_2')\,f_{2,2}^{[s]}[l_1])\,x_2] \tag{13.47}$$

Substituting equations 13.19, 13.20, 13.16, 13.46, and 13.47 in equations 13.29-32, 13.39, and 13.40 and separating Fourier orders, the following equations are obtained,

$$f_2^{[p,j]}[n_1,n_2]\,ffE_3^{[j]}[n_1,n_2] - f_3^{[p,j]}[n_1,n_2]\,ffE_2^{[j]}[n_1,n_2] = \tfrac{1}{\lambda}\,ffH_1^{[j]}[n_1,n_2] \tag{13.48}$$

$$i\,2\pi\,f_3^{[p,j]}[n_1,n_2]\,ffE_1^{[j]}[n_1,n_2] - \partial_1\,ffE_3^{[j]}[n_1,n_2] = i\tfrac{2\pi}{\lambda}\,ffH_2^{[j]}[n_1,n_2] \tag{13.49}$$

$$\partial_1\,ffE_2^{[j]}[n_1,n_2] - i\,2\pi\,f_2^{[p,j]}[n_1,n_2]\,ffE_1^{[j]}[n_1,n_2] = i\tfrac{2\pi}{\lambda}\,ffH_3^{[j]}[n_1,n_2] \tag{13.50}$$

$$f_2^{[p,j]}[n_1,n_2]\,ffH_3^{[j]}[n_1,n_2] - f_3^{[p,j]}[n_1,n_2]\,ffH_2^{[j]}[n_1,n_2] = -\tfrac{1}{\lambda}\sum_{n_1',n_2'} ff\varepsilon 1[l_1,n_1-n_1',n_2-n_2']\,ffE_1^{[j]}[n_1',n_2'] \tag{13.51}$$

$$\sum_{n_1',n_2'} frt\varepsilon 1[l_1,n_1-n_1',n_2,n_2']\left(i\,2\pi\,f_3^{[p,j]}[n_1',n_2']\,ffH_1^{[j]}[n_1',n_2'] - \partial_1\,ffH_3^{[j]}[n_1',n_2']\right) = -i\tfrac{2\pi}{\lambda}\,ffE_2^{[j]}[n_1,n_2] \tag{13.52}$$

$$\partial_1\,ffH_2^{[j]}[n_1,n_2] - i\,2\pi\,f_2^{[p,j]}[n_1,n_2]\,ffH_1^{[j]}[n_1,n_2] = -i\tfrac{2\pi}{\lambda}\sum_{n_1',n_2'} frtr\varepsilon 1[l_1,n_1-n_1',n_2,n_2']\,ffE_3^{[j]}[n_1',n_2'] \tag{13.53}$$

Two modifications will be made in the above equations to recast them in numerically tractable form. The first modification, "index truncation", restricts the summation indices to a finite range, and the second modification, “index enumeration”, restates the equations using simplified matrix notation.

The set of order index pairs (m_1, m_2) retained in the field expansions 5.7 and 5.12 is defined by the limit conditions $m_2 \in \mathcal{M}_2$, $m_1 \in \mathcal{M}_1[m_2]$ (conditions 4.11 and 4.12). This set is partitioned into decoupled index blocks corresponding to the base index pairs $(m_1^{[0]}, m_2^{[0]})$ in equation 13.4, and the set of suborder indices (n_1, n_2) corresponding to block j is represented by conditions analogous to 4.11 and 4.12

$$n_2 \in \mathcal{N}_2^{[j]} \tag{13.54}$$

$$n_1 \in \mathcal{N}_1^{[j]}[n_2] \tag{13.55}$$

The summation indices n_1 and n_2 in equations 13.12, 13.13, and 13.17-20 will be restricted to this range. Also, the index n_2' in equations 13.32-34 and 13.37-40 will be limited to the range $n_2' \in \mathcal{N}_2^{[j]}$. (Note that this may make the definitions of $rt\varepsilon 1$ and $rtr\varepsilon 1$ implicitly dependent on j in equations 13.37 and 13.38.) The same limit applies in equations 13.51-53, and the index n_1' is also limited to $n_1' \in \mathcal{N}_1^{[j]}[n_2']$.

The equations are further simplified by applying the index enumeration 9.1 to the suborder indices (n_1, n_2) in each index block,

$$\begin{aligned}
&(n_1, n_2) = (n_1^{[\text{enum}]}[j,k], n_2^{[\text{enum}]}[j,k]) \\
&(\text{and similarly, } (n_1', n_2') = (n_1^{[\text{enum}]}[j,k'], n_2^{[\text{enum}]}[j,k'])), \\
&\quad j = 1 \ldots j^{[\max]}, \quad k = 1 \ldots k^{[\max]}[j]
\end{aligned} \tag{13.56}$$

wherein $n_1^{[\text{enum}]}$ and $n_2^{[\text{enum}]}$ are defined from $m_1^{[\text{enum}]}$ and $m_2^{[\text{enum}]}$ in accordance with equation 13.4,

$$\begin{aligned}
&\left(m_1^{[\text{enum}]}[j,k] \quad m_2^{[\text{enum}]}[j,k] \right) = \left(m_1^{[0]}[j] \quad m_2^{[0]}[j] \right) + \\
&\quad \left(n_1^{[\text{enum}]}[j,k] \quad n_2^{[\text{enum}]}[j,k] \right) \begin{pmatrix} h_{1,1}[l_1] & h_{1,2}[l_1] \\ h_{2,1}[l_1] & h_{2,2}[l_1] \end{pmatrix}
\end{aligned} \tag{13.57}$$

The permittivity terms in equations 13.51-53 will be represented by Toeplitz matrices defined as follows,

$$tt\varepsilon 1[l_1, n_1, n_1', n_2, n_2'] = ff\varepsilon 1[l_1, n_1 - n_1', n_2 - n_2'] \tag{13.58}$$

$$trt\varepsilon 1[l_1, n_1, n_1', n_2, n_2'] = frt\varepsilon 1[l_1, n_1 - n_1', n_2, n_2'] \tag{13.59}$$

$$trtr\varepsilon 1[l_1, n_1, n_1', n_2, n_2'] = frtr\varepsilon 1[l_1, n_1 - n_1', n_2, n_2'] \tag{13.60}$$

The following notational abbreviations will be used for these matrices,

$$\begin{aligned} &tt\varepsilon 1_{k,k'} = \\ &\quad tt\varepsilon 1[l_1, n_1^{[\text{enum}]}[j,k], n_1^{[\text{enum}]}[j,k'], n_2^{[\text{enum}]}[j,k], n_2^{[\text{enum}]}[j,k']] \end{aligned} \tag{13.61}$$

$$\begin{aligned} &trt\varepsilon 1_{k,k'} = \\ &\quad trt\varepsilon 1[l_1, n_1^{[\text{enum}]}[j,k], n_1^{[\text{enum}]}[j,k'], n_2^{[\text{enum}]}[j,k], n_2^{[\text{enum}]}[j,k']] \end{aligned} \tag{13.62}$$

$$\begin{aligned} &trtr\varepsilon 1_{k,k'} = \\ &\quad trtr\varepsilon 1[l_1, n_1^{[\text{enum}]}[j,k], n_1^{[\text{enum}]}[j,k'], n_2^{[\text{enum}]}[j,k], n_2^{[\text{enum}]}[j,k']] \end{aligned} \tag{13.63}$$

(The matrices' dependence on l_1 and j is implicit.) Similar shorthand will be used for the field quantities in equations 13.48-53,

$$\begin{aligned} &ffE_{\nu,k}^{[j]} = ffE_\nu^{[j]}[n_1^{[\text{enum}]}[j,k], n_2^{[\text{enum}]}[j,k]] = ffE_\nu[m_1^{[\text{enum}]}[j,k], m_2^{[\text{enum}]}[j,k]]; \\ &\quad \nu = 1,2,3 \end{aligned} \tag{13.64}$$

$$\begin{aligned} &ffH_{\nu,k}^{[j]} = ffH_\nu^{[j]}[n_1^{[\text{enum}]}[j,k], n_2^{[\text{enum}]}[j,k]] = ffH_\nu[m_1^{[\text{enum}]}[j,k], m_2^{[\text{enum}]}[j,k]]; \\ &\quad \nu = 1,2,3 \end{aligned} \tag{13.65}$$

(cf. definitions 13.8 and 13.9, and equation 13.57). The tangential frequency terms will be represented as

$$f_{\nu,k}^{[p,j]} = f_\nu^{[p,j]}[n_1^{[\text{enum}]}[j,k], n_2^{[\text{enum}]}[j,k]]; \quad \nu = 2,3 \tag{13.66}$$

With the above substitutions, equations 13.48-51 simplify to

$$f_{2,k}^{[p,j]}\, ffE_{3,k}^{[j]} - f_{3,k}^{[p,j]}\, ffE_{2,k}^{[j]} = \tfrac{1}{\lambda}\, ffH_{1,k}^{[j]} \tag{13.67}$$

$$i\,2\pi\, f_{3,k}^{[p,j]}\, ffE_{1,k}^{[j]} - \partial_1\, ffE_{3,k}^{[j]} = i\,\tfrac{2\pi}{\lambda}\, ffH_{2,k}^{[j]} \tag{13.68}$$

$$\partial_1\, ffE_{2,k}^{[j]} - i\,2\pi\, f_{2,k}^{[p,j]}\, ffE_{1,k}^{[j]} = i\,\tfrac{2\pi}{\lambda}\, ffH_{3,k}^{[j]} \tag{13.69}$$

$$f_{2,k}^{[p,j]}\, ffH_{3,k}^{[j]} - f_{3,k}^{[p,j]}\, ffH_{2,k}^{[j]} = -\tfrac{1}{\lambda}\sum_{k'} tt\varepsilon 1_{k,k'}\, ffE_{1,k'}^{[j]} \tag{13.70}$$

$$\sum_{k'} trt\varepsilon 1_{k,k'} \left(i\,2\pi\, f_{3,k'}^{[p,j]}\, ffH_{1,k'}^{[j]} - \partial_1\, ffH_{3,k'}^{[j]} \right) = -i\,\tfrac{2\pi}{\lambda}\, ffE_{2,k}^{[j]} \tag{13.71}$$

$$\partial_1\, ffH_{2,k}^{[j]} - i\,2\pi\, f_{2,k}^{[p,j]}\, ffH_{1,k}^{[j]} = -i\,\tfrac{2\pi}{\lambda} \sum_{k'} trtr\varepsilon 1_{k,k'}\, ffE_{3,k'}^{[j]} \tag{13.72}$$

As a notational convenience, the frequency factors in the above equations will be represented as diagonal matrices,

$$df_{\nu,k,k'}^{[p,j]} = \begin{cases} f_{\nu,k}^{[p,j]} & \text{if } k = k' \\ 0 & \text{if } k \neq k' \end{cases}; \quad \nu = 2,3 \tag{13.73}$$

Leaving out the k and k' indices, equations 13.67-72 can be expressed more compactly in matrix form,

$$df_2^{[p,j]}\, ffE_3^{[j]} - df_3^{[p,j]}\, ffE_2^{[j]} = \tfrac{1}{\lambda}\, ffH_1^{[j]} \tag{13.74}$$

$$i\,2\pi\, df_3^{[p,j]}\, ffE_1^{[j]} - \partial_1\, ffE_3^{[j]} = i\,\tfrac{2\pi}{\lambda}\, ffH_2^{[j]} \tag{13.75}$$

$$\partial_1\, ffE_2^{[j]} - i\,2\pi\, df_2^{[p,j]}\, ffE_1^{[j]} = i\,\tfrac{2\pi}{\lambda}\, ffH_3^{[j]} \tag{13.76}$$

$$df_2^{[p,j]}\, ffH_3^{[j]} - df_3^{[p,j]}\, ffH_2^{[j]} = -\tfrac{1}{\lambda}\, tt\varepsilon 1\, ffE_1^{[j]} \tag{13.77}$$

$$trt\varepsilon 1 \left(i\,2\pi\, df_3^{[p,j]}\, ffH_1^{[j]} - \partial_1\, ffH_3^{[j]} \right) = -i\,\tfrac{2\pi}{\lambda}\, ffE_2^{[j]} \tag{13.78}$$

$$\partial_1\, ffH_2^{[j]} - i\,2\pi\, df_2^{[p,j]}\, ffH_1^{[j]} = -i\,\tfrac{2\pi}{\lambda}\, trtr\varepsilon 1\, ffE_3^{[j]} \tag{13.79}$$

The permittivity terms in equations 13.77 and 13.78 are moved to the other side of each equation,

$$rtt\varepsilon 1 \left(df_2^{[p,j]}\, ffH_3^{[j]} - df_3^{[p,j]}\, ffH_2^{[j]} \right) = -\tfrac{1}{\lambda}\, ffE_1^{[j]} \tag{13.80}$$

$$i\,2\pi\, df_3^{[p,j]}\, ffH_1^{[j]} - \partial_1\, ffH_3^{[j]} = -i\,\tfrac{2\pi}{\lambda}\, rtrt\varepsilon 1\, ffE_2^{[j]} \tag{13.81}$$

wherein

$$rtt\varepsilon 1 = (tt\varepsilon 1)^{-1} \tag{13.82}$$

$$rtrt\varepsilon 1 = (trt\varepsilon 1)^{-1} \tag{13.83}$$

Equations 13.74 and 13.80 are used to eliminate $ffE_1^{[j]}$ and $ffH_1^{[j]}$ from equations 13.75, 13.76, 13.81, and 13.79,

$$\partial_1\, ffE_3^{[j]} = -i\,2\pi\left(\tfrac{1}{\lambda}\, ffH_2^{[j]} + \lambda\, df_3^{[p,j]}\, rtt\varepsilon 1 (df_2^{[p,j]}\, ffH_3^{[j]} - df_3^{[p,j]}\, ffH_2^{[j]})\right) \tag{13.84}$$

$$\partial_1\, ffE_2^{[j]} = i\,2\pi\left(\tfrac{1}{\lambda}\, ffH_3^{[j]} - \lambda\, df_2^{[p,j]}\, rtt\varepsilon 1 (df_2^{[p,j]}\, ffH_3^{[j]} - df_3^{[p,j]}\, ffH_2^{[j]})\right) \tag{13.85}$$

$$\partial_1\, ffH_3^{[j]} = i\,2\pi\left(\tfrac{1}{\lambda}\, rtrt\varepsilon 1\, ffE_2^{[j]} + \lambda\, df_3^{[p,j]}\left(df_2^{[p,j]}\, ffE_3^{[j]} - df_3^{[p,j]}\, ffE_2^{[j]}\right)\right) \tag{13.86}$$

$$\partial_1\, ffH_2^{[j]} = -i\,2\pi\left(\tfrac{1}{\lambda}\, trtr\varepsilon 1\, ffE_3^{[j]} - \lambda\, df_2^{[p,j]}\left(df_2^{[p,j]}\, ffE_3^{[j]} - df_3^{[p,j]}\, ffE_2^{[j]}\right)\right) \tag{13.87}$$

Equations 13.84-87 will be restated in terms of the fields' $\hat{s}$ and $\hat{q}$ projections, and will then be separated into up and down waves relative to a fictitious homogeneous medium of permittivity $\varepsilon^{[f]}$, as described in section 6 (equations 6.50-53). The up/down separation involves the $f_1^{[\pm]}$ terms defined in equation 6.35 (with $\varepsilon^{[c]} = \varepsilon^{[f]}$), which will be specialized for index block j as follows,

$$\begin{aligned} f_{1,k}^{[\pm,j]} &= f_1^{[\pm]}[m_1^{[\mathrm{enum}]}[j,k], m_2^{[\mathrm{enum}]}[j,k]] \\ &= \pm\sqrt{\frac{\varepsilon^{[f]}}{\lambda^2} - (f_{2,k}^{[p,j]})^2 - (f_{3,k}^{[p,j]})^2} \end{aligned} \tag{13.88}$$

($\varepsilon^{[f]}$ should have a positive imaginary part to ensure that $f_{1,k}^{[\pm,j]} \neq 0$.) The grating-tangential $\hat{s}$ vector defined by equations 6.38 and 6.39 is similarly specialized for index block j,

$$\hat{s}_k^{[j]} = \hat{e}_2\, s_{2,k}^{[j]} + \hat{e}_3\, s_{3,k}^{[j]} = \hat{s}[m_1^{[\mathrm{enum}]}[j,k], m_2^{[\mathrm{enum}]}[j,k]] \tag{13.89}$$

wherein

$$\left(s_{2,k}^{[j]},\quad s_{3,k}^{[j]}\right) = \begin{cases} \dfrac{\left(-f_{3,k}^{[p,j]},\quad f_{2,k}^{[p,j]}\right)}{\sqrt{(f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2}} & \text{if } (f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2 \neq 0 \\ \dfrac{\left(-f_{3,1}^{[g]},\quad f_{2,1}^{[g]}\right)}{\sqrt{(f_{2,1}^{[g]})^2 + (f_{3,1}^{[g]})^2}} & \text{if } (f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2 = 0 \end{cases} \tag{13.90}$$

(The "if … = 0" branch of this definition may be modified to ensure polarization decoupling for uniperiodic gratings – see section 15 – but the modification has no effect on the results in this section.) The grating-tangential $\hat{q}$ vector (equation 6.41) is represented as

$$\begin{aligned} \hat{q}_k^{[\pm,j]} &= \hat{q}^{[\pm]}[m_1^{[\text{enum}]}[j,k], m_2^{[\text{enum}]}[j,k]] \\ &= \pm \hat{s}_k^{[j]} \times \hat{e}_1 = \pm(\hat{e}_2\, s_{3,k}^{[j]} - \hat{e}_3\, s_{2,k}^{[j]}) \end{aligned} \tag{13.91}$$

The following diagonal matrix notation will be used for $f_1^{[\pm,j]}$ and $s_\nu^{[j]}$,

$$df_{1,k,k'}^{[\pm,j]} = \begin{cases} f_{1,k}^{[\pm,j]} & \text{if } k = k' \\ 0 & \text{if } k \neq k' \end{cases} \tag{13.92}$$

$$ds_{\nu,k,k'}^{[j]} = \begin{cases} s_{\nu,k}^{[j]} & \text{if } k = k' \\ 0 & \text{if } k \neq k' \end{cases}; \quad \nu = 2,3 \tag{13.93}$$

Equations 6.50-53 are adapted to define the fields' $\hat{s}$ and $\hat{q}$ projections ($ffE_s^{[j]}$, $ffE_q^{[j]}$, $ffH_s^{[j]}$, $ffH_q^{[j]}$), and also define the corresponding up/down fields' $\hat{s}$ projections ($ffE_s^{[\pm,j]}$, $ffH_s^{[\pm,j]}$),

$$\begin{aligned} ffE_{s,k}^{[j]}[x_1] &= \hat{s}_k^{[j]} \bullet ff\vec{E}_k^{[j]}[x_1] = s_{2,k}^{[j]}\, ffE_{2,k}^{[j]}[x_1] + s_{3,k}^{[j]}\, ffE_{3,k}^{[j]}[x_1] \\ &= ffE_{s,k}^{[+,j]}[x_1] + ffE_{s,k}^{[-,j]}[x_1] \end{aligned} \tag{13.94}$$

$$\begin{aligned} ffH_{s,k}^{[j]}[x_1] &= \hat{s}_k^{[j]} \bullet ff\vec{H}_k^{[j]}[x_1] = s_{2,k}^{[j]}\, ffH_{2,k}^{[j]}[x_1] + s_{3,k}^{[j]}\, ffH_{3,k}^{[j]}[x_1] \\ &= ffH_{s,k}^{[+,j]}[x_1] - ffH_{s,k}^{[-,j]}[x_1] \end{aligned} \tag{13.95}$$

$$\begin{aligned} ffE_{q,k}^{[j]}[x_1] &= \hat{q}_k^{[+,j]} \bullet ff\vec{E}_k^{[j]}[x_1] = s_{3,k}^{[j]}\, ffE_{2,k}^{[j]}[x_1] - s_{2,k}^{[j]}\, ffE_{3,k}^{[j]}[x_1] \\ &= \frac{\lambda}{\varepsilon^{[f]}}\, f_{1,k}^{[+,j]}\, (ffH_{s,k}^{[+,j]}[x_1] + ffH_{s,k}^{[-,j]}[x_1]) \end{aligned} \tag{13.96}$$

$$\begin{aligned} ffH_{q,k}^{[j]}[x_1] &= \hat{q}_k^{[+,j]} \bullet ff\vec{H}_k^{[j]}[x_1] = s_{3,k}^{[j]}\, ffH_{2,k}^{[j]}[x_1] - s_{2,k}^{[j]}\, ffH_{3,k}^{[j]}[x_1] \\ &= -\lambda\, f_{1,k}^{[+,j]}\, (ffE_{s,k}^{(+,j)}[x_1] - ffE_{s,k}^{[-,j]}[x_1]) \end{aligned} \tag{13.97}$$

Using definitions 13.92 and 13.93, these equations can be stated more compactly as

$$\begin{aligned} ffE_s^{[j]} &= ds_2^{[j]}\, ffE_2^{[j]} + ds_3^{[j]}\, ffE_3^{[j]} \\ &= ffE_s^{[+,j]} + ffE_s^{[-,j]} \end{aligned} \tag{13.98}$$

$$\begin{aligned} ffH_s^{[j]} &= ds_2^{[j]}\, ffH_2^{[j]} + ds_3^{[j]}\, ffH_3^{[j]} \\ &= ffH_s^{[+,j]} - ffH_s^{[-,j]} \end{aligned} \tag{13.99}$$

$$\begin{aligned} ffE_q^{[j]} &= ds_3^{[j]}\, ffE_2^{[j]} - ds_2^{[j]}\, ffE_3^{[j]} \\ &= \frac{\lambda}{\varepsilon^{[f]}} df_1^{[+,j]} \, (ffH_s^{[+,j]} + ffH_s^{[-,j]}) \end{aligned} \tag{13.100}$$

$$\begin{aligned} ffH_q^{[j]} &= ds_3^{[j]}\, ffH_2^{[j]} - ds_2^{[j]}\, ffH_3^{[j]} \\ &= -\lambda\, df_1^{[+,j]} \, (ffE_s^{[+,j]} - ffE_s^{[-,j]}) \end{aligned} \tag{13.101}$$

The above equations are used to restate differential equations 13.84-87 in terms of the fields' $\hat{s}$ and $\hat{q}$ projections,

$$\partial_1\, ffE_s^{[j]} = -i\, \tfrac{2\pi}{\lambda}\, ffH_q^{[j]} \tag{13.102}$$

$$\begin{aligned} \partial_1\, ffE_q^{[j]} = i\, 2\pi \Big(&\tfrac{1}{\lambda}\, ffH_s^{[j]} \\ &- \lambda \sqrt{(df_2^{[p,j]})^2 + (df_3^{[p,j]})^2}\; rtt\varepsilon 1 \sqrt{(df_2^{[p,j]})^2 + (df_3^{[p,j]})^2}\; ffH_s^{[j]} \Big) \end{aligned} \tag{13.103}$$

$$\begin{aligned} \partial_1\, ffH_s^{[j]} = i\, \tfrac{2\pi}{\lambda} \Big(&ds_3^{[j]}\, rtrt\varepsilon 1 (ds_2^{[j]}\, ffE_s^{[j]} + ds_3^{[j]}\, ffE_q^{[j]}) \\ &- ds_2^{[j]}\, trtr\varepsilon 1 (ds_3^{[j]}\, ffE_s^{[j]} - ds_2^{[j]}\, ffE_q^{[j]}) \Big) \end{aligned} \tag{13.104}$$

$$\begin{aligned} \partial_1\, ffH_q^{[j]} = -i\, 2\pi \Big(&\tfrac{1}{\lambda} \Big(ds_2^{[j]}\, rtrt\varepsilon 1 (ds_2^{[j]}\, ffE_s^{[j]} + ds_3^{[j]}\, ffE_q^{[j]}) \\ &\quad + ds_3^{[j]}\, trtr\varepsilon 1 (ds_3^{[j]}\, ffE_s^{[j]} - ds_2^{[j]}\, ffE_q^{[j]}) \Big) \\ &- \lambda \Big((f_2^{[p,j]})^2 + (f_3^{[p,j]})^2 \Big) ffE_s^{[j]} \Big) \end{aligned} \tag{13.105}$$

These equations are of the form

$$\partial_1 \begin{pmatrix} ffE_s^{[j]} \\ ffE_q^{[j]} \end{pmatrix} = i\, DEH \begin{pmatrix} ffH_s^{[j]} \\ ffH_q^{[j]} \end{pmatrix} \tag{13.106}$$

$$\partial_1 \begin{pmatrix} ffH_s^{[j]} \\ ffH_q^{[j]} \end{pmatrix} = i\, DHE \begin{pmatrix} ffE_s^{[j]} \\ ffE_q^{[j]} \end{pmatrix} \tag{13.107}$$

wherein

$$DEH = 2\pi \begin{pmatrix} \boldsymbol{0} & -\frac{1}{\lambda}\boldsymbol{I} \\ \frac{1}{\lambda}\boldsymbol{I} - \lambda\sqrt{(df_2^{[p,j]})^2 + (df_3^{[p,j]})^2}\; rtt\varepsilon 1 \sqrt{(df_2^{[p,j]})^2 + (df_3^{[p,j]})^2} & \boldsymbol{0} \end{pmatrix}$$

(13.108)

$$DHE = 2\pi \left(\frac{1}{\lambda} \left(\begin{pmatrix} ds_3^{[j]} \\ -ds_2^{[j]} \end{pmatrix} rtrt\varepsilon 1 \begin{pmatrix} ds_2^{[j]} & ds_3^{[j]} \end{pmatrix} - \begin{pmatrix} ds_2^{[j]} \\ ds_3^{[j]} \end{pmatrix} trtr\varepsilon 1 \begin{pmatrix} ds_3^{[j]} & -ds_2^{[j]} \end{pmatrix} \right) \right.$$
$$\left. + \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \lambda\left((df_2^{[p,j]})^2 + (df_3^{[p,j]})^2\right) & \boldsymbol{0} \end{pmatrix} \right)$$

(13.109)

Given initial values of the field quantities at position $x_1 = x_1^{[0]}$ in the stratum, their values at any other position $x_1 = x_1^{[0]} + \Delta x_1$ in the stratum are given by

$$\begin{pmatrix} ffE_s^{[j]}[x_1^{[0]} + \Delta x_1] \\ ffE_q^{[j]}[x_1^{[0]} + \Delta x_1] \end{pmatrix} = \Phi EE \begin{pmatrix} ffE_s^{[j]}[x_1^{[0]}] \\ ffE_q^{[j]}[x_1^{[0]}] \end{pmatrix} + i\,\Phi EH \begin{pmatrix} ffH_s^{[j]}[x_1^{[0]}] \\ ffH_q^{[j]}[x_1^{[0]}] \end{pmatrix} \tag{13.110}$$

$$\begin{pmatrix} ffH_s^{[j]}[x_1^{[0]} + \Delta x_1] \\ ffH_q^{[j]}[x_1^{[0]} + \Delta x_1] \end{pmatrix} = -i\,\Phi HE \begin{pmatrix} ffE_s^{[j]}[x_1^{[0]}] \\ ffE_q^{[j]}[x_1^{[0]}] \end{pmatrix} + \Phi HH \begin{pmatrix} ffH_s^{[j]}[x_1^{[0]}] \\ ffH_q^{[j]}[x_1^{[0]}] \end{pmatrix} \tag{13.111}$$

wherein the Φ 's have the form of an exponential matrix,

$$\begin{pmatrix} \Phi EE & \Phi EH \\ \Phi HE & \Phi HH \end{pmatrix} = \exp\left[\begin{pmatrix} \boldsymbol{0} & DEH \\ -DHE & \boldsymbol{0} \end{pmatrix} \Delta x_1 \right] \tag{13.112}$$

If Δx_1 is set equal to $x_1^{[1]} - x_1^{[0]}$, then the Φ matrix on the left side of equation 13.112 will represent a linear relationship between the field quantities at $x_1 = x_1^{[0]}$ and $x_1 = x_1^{[1]}$, from which the stratum's S matrix can be determined. However, the exponential matrix computation can become numerically unstable when Δx_1 is large [Ref. 5], so an alternative approach will be used. Δx_1 is set equal to the stratum thickness scaled by a power of 1/2,

$$\Delta x_1 = (x_1^{[1]} - x_1^{[0]}) / 2^{sp} \tag{13.113}$$

wherein the scaling factor's exponent sp ("scaling power") is chosen so that the exponential matrix can be calculated using a rational approximation. (Details of the calculation and the choice of sp are outlined in Appendix B.) The exponential matrix is then converted into an S matrix representing a stratum of thickness Δx_1, and sp stacking operations are performed (each one doubling the thickness) to obtain the S matrix for a stratum of thickness $x_1^{[1]} - x_1^{[0]}$.

The i coefficient in equations 13.106, 13.107, 13.110 and 13.111 is factored out of the exponential matrix so that the matrix is real-valued for a symmetric, non-absorbing stratum, i.e.,

$$\begin{aligned}&\varepsilon 1[l_1][x_2, x_3] = \varepsilon 1[l_1][-x_2, -x_3] \quad , \mathrm{Im}[\varepsilon 1[l_1][x_2, x_3]] = 0 \quad \rightarrow \\ &\quad \mathrm{Im}[DEH] = 0,\ \mathrm{Im}[DHE] = 0, \\ &\quad \mathrm{Im}[\Phi EE] = 0,\ \mathrm{Im}[\Phi EH] = 0,\ \mathrm{Im}[\Phi HE] = 0,\ \mathrm{Im}[\Phi HH] = 0\end{aligned} \tag{13.114}$$

In some cases, a coordinate translation can be applied, as described in section 12, to move the (x_2, x_3) coordinate origin to the symmetry axis so that the condition $\varepsilon 1[l_1][x_2, x_3] = \varepsilon 1[l_1][-x_2, -x_3]$ holds.

Equations 13.98-101 can be used to restate equations 13.110 and 13.111 in terms of the up/down field amplitudes, $ffE_s^{[\pm, j]}$ and $ffH_s^{[\pm, j]}$. Defining

$$F^{[\pm, j]} = \begin{pmatrix} ffE_s^{[\pm, j]} \\ ffH_s^{[\pm, j]} \end{pmatrix} \tag{13.115}$$

equations 13.110 and 13.111 translate to

$$\begin{pmatrix} F^{[+]}[x_1^{[0]} + \Delta x_1] \\ F^{[-]}[x_1^{[0]} + \Delta x_1] \end{pmatrix} = \begin{pmatrix} \Phi^{[++]} & \Phi^{[+-]} \\ \Phi^{[-+]} & \Phi^{[--]} \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[0]}] \end{pmatrix} \tag{13.116}$$

wherein

$$\begin{aligned}&\Phi^{[\pm\pm']} = \\ &\frac{1}{2}\left(\begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \varepsilon^{[f]} (\lambda\, df_1^{[+,j]})^{-1} \end{pmatrix} \left(\Phi EE \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & (\lambda / \varepsilon^{[f]})\, df_1^{[+,j]} \end{pmatrix} \pm' i\, \Phi EH \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\lambda\, df_1^{[+,j]} & \boldsymbol{0} \end{pmatrix} \right) \right. \\ &\left. \quad \pm \begin{pmatrix} \boldsymbol{0} & -(\lambda\, df_1^{[+,j]})^{-1} \\ \boldsymbol{I} & \boldsymbol{0} \end{pmatrix} \left(-i\, \Phi HE \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & (\lambda / \varepsilon^{[f]})\, df_1^{[+,j]} \end{pmatrix} \pm' \Phi HH \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\lambda\, df_1^{[+,j]} & \boldsymbol{0} \end{pmatrix} \right) \right)\end{aligned} \tag{13.117}$$

In equation 13.117 the $\pm$ signs (unprimed) are correlated, and the $\pm'$ signs (primed) are correlated. The $df_1^{[+,j]}$ term is guaranteed to be non-singular if $\varepsilon^{[f]}$ has a non-zero imaginary part (see equation 13.88).

Equation 13.116 can be rearranged as follows, with the waves entering the stratum on the right side of the equation and the exiting waves on the left,

$$\begin{pmatrix} -\Phi^{[+-]} & \boldsymbol{I} \\ \Phi^{[--]} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} F^{[-]}[x_1^{[0]}] \\ F^{[+]}[x_1^{[0]} + \Delta x_1] \end{pmatrix} = \begin{pmatrix} \Phi^{[++]} & \boldsymbol{0} \\ -\Phi^{[-+]} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[0]} + \Delta x_1] \end{pmatrix} \quad (13.118)$$

This equations can be solved for the exiting waves to determine the S matrix,

$$\begin{pmatrix} F^{[-]}[x_1^{[0]}] \\ F^{[+]}[x_1^{[0]} + \Delta x_1] \end{pmatrix} = \begin{pmatrix} S00 & S01 \\ S10 & S11 \end{pmatrix} \begin{pmatrix} F^{[+]}[x_1^{[0]}] \\ F^{[-]}[x_1^{[0]} + \Delta x_1] \end{pmatrix} \quad (13.119)$$

wherein

$$\begin{aligned} \begin{pmatrix} S00 & S01 \\ S10 & S11 \end{pmatrix} &= \begin{pmatrix} \boldsymbol{0} & (\Phi^{[--]})^{-1} \\ \boldsymbol{I} & \Phi^{[+-]}(\Phi^{[--]})^{-1} \end{pmatrix} \begin{pmatrix} \Phi^{[++]} & \boldsymbol{0} \\ -\Phi^{[-+]} & \boldsymbol{I} \end{pmatrix} \\ &= \begin{pmatrix} -(\Phi^{[--]})^{-1}\Phi^{[-+]} & (\Phi^{[--]})^{-1} \\ \Phi^{[++]} - \Phi^{[+-]}(\Phi^{[--]})^{-1}\Phi^{[-+]} & \Phi^{[+-]}(\Phi^{[--]})^{-1} \end{pmatrix} \end{aligned} \quad (13.120)$$

It is not necessary to compute the entire S matrix. Based on the symmetry relations 7.3 ($S00 = S11, S10 = S01$) only $S01$ and $S11$ need be determined. Taking advantage of this symmetry, the following stacking operation effectively doubles Δx_1,

$$\begin{pmatrix} S01 \\ S11 \end{pmatrix} \leftarrow \begin{pmatrix} S01(\boldsymbol{I} - S11^2)^{-1} S01 \\ S11 + S01\, S11(\boldsymbol{I} - S11^2)^{-1} S01 \end{pmatrix} \qquad (\Delta x_1 \leftarrow 2\,\Delta x_1) \qquad (13.121)$$

(from equations 7.3 and 8.8, with $Sa = Sb = S$). This operation is repeated sp times to build up the full S matrix for a stratum of thickness $x_1^{[1]} - x_1^{[0]} = 2^{sp}\,\Delta x_1$ (cf. equation 13.113).

14. The S matrix for a uniperiodic stratum

A uniperiodic stratum can be treated as a special case of a biperiodic stratum, requiring just a few specializations or modifications of the equations in section 13. In equation 13.1 the range of n_2 is limited to $\{0\}$,

$$\varepsilon[\vec{x}] = \varepsilon 1[l_1][x_2, x_3] = \sum_{n_1} \mathit{ff}\varepsilon 1[l_1, n_1, 0] \exp[i\, 2\pi\, n_1\, \vec{f}_1^{[s]}[l_1] \bullet \vec{x}] \qquad (14.1)$$

Equation 13.4 reduces to

$$\left(m_1, \quad m_2\right) = \left(m_1^{[0]}, \quad m_2^{[0]}\right) + n_1 \left(h_{1,1}[l_1], \quad h_{1,2}[l_1]\right) \tag{14.2}$$

and equation 13.5 is replaced with the following definition of n_1

$$\left.\begin{array}{l} \text{If } \left|h_{1,1}[l_1]\right| \ge \left|h_{1,2}[l_1]\right|: \quad n_1 = \text{floor}[m_1 / h_{1,1}[l_1]] \\ \text{If } \left|h_{1,1}[l_1]\right| < \left|h_{1,2}[l_1]\right|: \quad n_1 = \text{floor}[m_2 / h_{1,2}[l_1]] \end{array}\right\} \tag{14.3}$$

($h_{1,1}[l_1]$ and $h_{1,2}[l_1]$ cannot both be zero; see condition 3.21.) Relations 13.6 and 13.7 are replaced by a single relation constraining just one of the base indices $m_1^{[0]}$ or $m_2^{[0]}$,

$$\left.\begin{array}{l} \text{If } \left|h_{1,1}[l_1]\right| \ge \left|h_{1,2}[l_1]\right|: \quad 0 \le m_1^{[0]} / h_{1,1}[l_1] < 1 \\ \text{If } \left|h_{1,1}[l_1]\right| < \left|h_{1,2}[l_1]\right|: \quad 0 \le m_2^{[0]} / h_{1,2}[l_1] < 1 \end{array}\right\} \tag{14.4}$$

Since the other base index is unconstrained, there is an infinite number of base index pairs $(m_1^{[0]}, m_2^{[0]})$ and associated decoupled index blocks. But in practice the order truncation (conditions 4.11 and 4.12) will eliminate all but a finite number of index blocks.

All remaining equations containing n_2 are modified according to the restriction $n_2 = 0$. In particular, equation 13.15 and 13.28 reduce to

$$\varepsilon[\vec{x}] = \varepsilon 1[l_1][x_2, x_3] = f\varepsilon 1[l_1, 0][x_2] \tag{14.5}$$

$$\frac{1}{\varepsilon[\vec{x}]} = \frac{1}{\varepsilon 1[l_1][x_2, x_3]} = fr\varepsilon 1[l_1, 0][x_2] \tag{14.6}$$

(i.e. the coordinate orientation defined by equations 13.14 makes the stratum permittivity independent of x_3.) The periodicity conditions 13.41 and 13.42 are replaced with a single condition,

$$\varepsilon 1[l_1][x_2 + 1 / f_{2,1}^{[s]}[l_1], x_3] = \varepsilon 1[l_1][x_2, x_3] \tag{14.7}$$

Substituting equation 13.14 and $n_2 = 0$ in equation 13.11 yields

$$\vec{f}^{[p,j]}[n_1, n_2] = \vec{f}^{[ip,j]} + n_1 \, \hat{e}_2 \, f_{2,1}^{[s]}[l_1] \tag{14.8}$$

Hence,

$$f_3^{[p,j]}[n_1, n_2] = f_3^{[ip,j]} \tag{14.9}$$

Thus, with the choice of coordinate orientation defined by equation 13.14, all of the field's spatial frequencies share the same $\hat{e}_3$ component. If $f_3^{[ip,j]} = 0$, these components will all be zero; and applying definition 13.66, the following conditions is obtained

$$f_3^{[ip,j]} = 0 \quad \rightarrow \quad f_{3,k}^{[p,j]} = 0 \tag{14.10}$$

Based on definition 13.90, it also follows that if $f_3^{[ip,j]} = 0$, then $s_{2,k}^{[j]}$ will be zero, with the possible exception of the case $(f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2 = 0$ and $f_{3,1}^{[g]} \neq 0$. (Section 15 describes a modified definition of $\hat{s}_k^{[j]}$ that can be used to avoid the exceptional condition, ensuring that $s_{2,k}^{[j]} = 0$ for all k when $f_3^{[ip,j]} = 0$.)

If $s_{2,k}^{[j]} = 0$ for all k, then differential equations 13.106 and 13.107 separate into decoupled TE ("Transverse Electric") and TM ("Transverse Magnetic") fields,

$$s_2^{[j]} = \boldsymbol{0} \quad \rightarrow \begin{cases} \partial_1\, ffE_s^{[j]} = i\, DEH^{[\text{TE}]}\, ffH_q^{[j]}, \quad \partial_1\, ffH_q^{[j]} = i\, DHE^{[\text{TE}]}\, ffE_s^{[j]} \quad (\text{TE}) \\ \partial_1\, ffE_q^{[j]} = i\, DEH^{[\text{TM}]}\, ffH_s^{[j]}, \quad \partial_1\, ffH_s^{[j]} = i\, DHE^{[\text{TM}]}\, ffE_q^{[j]} \quad (\text{TM}) \end{cases} \tag{14.11}$$

wherein

$$DEH^{[\text{TE}]} = -\tfrac{2\pi}{\lambda}\boldsymbol{I} \tag{14.12}$$

$$DHE^{[\text{TE}]} = 2\pi\left(-\tfrac{1}{\lambda}\, ds_3^{[j]}\, trtr\varepsilon 1\, ds_3^{[j]} + \lambda\left((df_2^{[p,j]})^2 + (df_3^{[p,j]})^2\right)\right) \tag{14.13}$$

$$DEH^{[\text{TM}]} = 2\pi\left(\tfrac{1}{\lambda}\boldsymbol{I} - \lambda\sqrt{(df_2^{[p,j]})^2 + (df_3^{[p,j]})^2}\; rtt\varepsilon 1 \sqrt{(df_2^{[p,j]})^2 + (df_3^{[p,j]})^2}\right) \tag{14.14}$$

$$DHE^{[\text{TM}]} = \tfrac{2\pi}{\lambda}\, ds_3^{[j]}\, rtrt\varepsilon 1\, ds_3^{[j]} \tag{14.15}$$

In this case, the matrices DEH and DHE (equations 13.108 and 13.109) have the following form,

$$DEH = \begin{pmatrix} \boldsymbol{0} & DEH^{[TE]} \\ DEH^{[TM]} & \boldsymbol{0} \end{pmatrix} \tag{14.16}$$

$$DHE = \begin{pmatrix} \boldsymbol{0} & DHE^{[TM]} \\ DHE^{[TE]} & \boldsymbol{0} \end{pmatrix} \tag{14.17}$$

Solutions to equations 14.11 are of the form

$$ffE_s^{[j]}[x_1^{[0]}+\Delta x_1] = \Phi EE^{[\mathrm{TE}]}\ ffE_s^{[j]}[x_1^{[0]}] + i\,\Phi EH^{[\mathrm{TE}]}\ ffH_q^{[j]}[x_1^{[0]}] \tag{14.18}$$

$$ffH_q^{[j]}[x_1^{[0]}+\Delta x_1] = -i\,\Phi HE^{[\mathrm{TE}]}\ ffE_s^{[j]}[x_1^{[0]}] + \Phi HH^{[\mathrm{TE}]}\ ffH_q^{[j]}[x_1^{[0]}] \tag{14.19}$$

$$ffE_q^{[j]}[x_1^{[0]}+\Delta x_1] = \Phi EE^{[\mathrm{TM}]}\ ffE_q^{[j]}[x_1^{[0]}] + i\,\Phi EH^{[\mathrm{TM}]}\ ffH_s^{[j]}[x_1^{[0]}] \tag{14.20}$$

$$ffH_s^{[j]}[x_1^{[0]}+\Delta x_1] = -i\,\Phi HE^{[\mathrm{TM}]}\ ffE_q^{[j]}[x_1^{[0]}] + \Phi HH^{[\mathrm{TM}]}\ ffH_s^{[j]}[x_1^{[0]}] \tag{14.21}$$

wherein

$$\begin{pmatrix} \Phi EE^{[\mathrm{TE}]} & \Phi EH^{[\mathrm{TE}]} \\ \Phi HE^{[\mathrm{TE}]} & \Phi HH^{[\mathrm{TE}]} \end{pmatrix} = \exp\left[\begin{pmatrix} \boldsymbol{0} & DEH^{[\mathrm{TE}]} \\ -DHE^{[\mathrm{TE}]} & \boldsymbol{0} \end{pmatrix} \Delta x_1\right] \tag{14.22}$$

$$\begin{pmatrix} \Phi EE^{[\mathrm{TM}]} & \Phi EH^{[\mathrm{TM}]} \\ \Phi HE^{[\mathrm{TM}]} & \Phi HH^{[\mathrm{TM}]} \end{pmatrix} = \exp\left[\begin{pmatrix} \boldsymbol{0} & DEH^{[\mathrm{TM}]} \\ -DHE^{[\mathrm{TM}]} & \boldsymbol{0} \end{pmatrix} \Delta x_1\right] \tag{14.23}$$

(Δx_1 is defined as in equation 13.113 with the scaling factor of 2^{-sp} applied, and sp stacking operations are subsequently applied to remove the scaling factor.)

With polarization decoupling, the index block corresponding to base indices $(m_1^{[0]}, m_2^{[0]})$, as defined in section 13, is split into two decoupled index blocks, both corresponding to the same base index pair $(m_1^{[0]}, m_2^{[0]})$ but corresponding to different polarizations (TE or TM). With this separation, each index block j has polarization indices $P^{[\mathrm{enum}]}[j,k]$ (equation 9.2) that are all identical (0 for TE, or 1 for TM),

$$P^{[\mathrm{enum}]}[j,k] = \begin{cases} 0 & \text{for TE} \\ 1 & \text{for TM} \end{cases} \tag{14.24}$$

Definition 13.115 is modified as followed for with polarization decoupling,

$$F^{[\pm,j]} = \begin{cases} ffE_s^{[\pm,j]} & \text{for TE} \\ ffH_s^{[\pm,j]} & \text{for TM} \end{cases} \tag{14.25}$$

The up/down field amplitudes $F^{[\pm,j]}$ are defined by equations 13.98-101. Based on these definitions, equations 14.18-21 translate to an equation having the same form as equation 13.116, wherein

$$\Phi^{[\pm\pm']} = \begin{cases} \dfrac{1}{2}\left(\begin{array}{l} \Phi EE^{[\mathrm{TE}]} \mp' i\,\Phi EH^{[\mathrm{TE}]}\,\lambda\, df_1^{[+,j]} \\ \quad \pm (\lambda\, df_1^{[+,j]})^{-1} \left(i\,\Phi HE^{[\mathrm{TE}]} \pm' \Phi HH^{[\mathrm{TE}]}\,\lambda\, df_1^{[+,j]} \right) \end{array} \right) & \text{for TE} \\ \dfrac{1}{2}\left(\begin{array}{l} \varepsilon^{[f]}\,(\lambda\, df_1^{[+,j]})^{-1} \left(\Phi EE^{[\mathrm{TM}]}\,(\lambda/\varepsilon^{[f]})\, df_1^{[+,j]} \pm' i\,\Phi EH^{[\mathrm{TM}]} \right) \\ \quad \pm \left(-i\,\Phi HE^{[\mathrm{TM}]}\,(\lambda/\varepsilon^{[f]})\, df_1^{[+,j]} \pm' \Phi HH^{[\mathrm{TM}]} \right) \end{array} \right) & \text{for TM} \end{cases} \tag{14.26}$$

The S matrix is then defined as in equation 13.120.

15. Change of $\hat{s}$ basis to maintain polarization decoupling

As noted in section 14, if $f_3^{[ip,j]} = 0$ then $s_{2,k}^{[j]}$ will be zero for all k, with the possible exception of the case $(f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2 = 0$ (see equation 13.90 and condition 14.10). However, the second branch condition in equation 13.90 ("if … = 0") can be modified to avoid the exceptional condition, ensuring polarization decoupling for all diffraction orders in the j-th index block when $f_3^{[ip,j]} = 0$. The second branch condition is modified to make $\hat{s}_k^{[j]}$ orthogonal to the stratum's basis frequency vector $\vec{f}_1^{[s]}[l_1]$, rather than the grating's basis frequency $\vec{f}_1^{[g]}$,

$$\left(s_{2,k}^{[j]},\quad s_{3,k}^{[j]} \right) = \begin{cases} \dfrac{\left(-f_{3,k}^{[p,j]},\quad f_{2,k}^{[p,j]} \right)}{\sqrt{(f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2}} & \text{if } (f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2 \neq 0 \\ \dfrac{\left(-f_{3,1}^{[s]}[l_1],\quad f_{2,1}^{[s]}[l_1] \right)}{\sqrt{(f_{2,1}^{[s]}[l_1])^2 + (f_{3,1}^{[s]}[l_1])^2}} & \text{if } (f_{2,k}^{[p,j]})^2 + (f_{3,k}^{[p,j]})^2 = 0 \end{cases} \tag{15.1}$$

With the coordinate orientation implicit in equations 13.14, the $f_{3,1}^{[s]}[l_1]$ term in equation 15.1 is zero, so that $s_{2,k}^{[j]}$ is zero.

None of the results of section 13 or 14 is affected by the adoption of definition 15.1 in lieu of 13.90; however, the S matrices for all strata must be modified to revert to the common definition 13.90 before stacking them to form the grating's cumulative S matrix. The $\hat{s}$ basis redefinition can itself be represented formally by an S matrix by the procedure outlined below.

No more than one of the tangential frequencies $\vec{f}^{[p]}[m_1, m_2]$ (definition 5.8) can be zero because the grating basis frequencies $\vec{f}_1^{[g]}$ and $\vec{f}_2^{[g]}$ are linearly independent (cf.

equations 3.25-27). We assume that $\vec{f}^{[p]}[m_1,m_2]$ is zero for particular order indices $(m_1,m_2)=(m_1^{[\perp]},m_2^{[\perp]})$,

$$(f_2^{[p]}[m_1^{[\perp]},m_2^{[\perp]}])^2+(f_3^{[p]}[m_1^{[\perp]},m_2^{[\perp]}])^2=0 \tag{15.2}$$

(The connotation of the "$\perp$" superscript is that the order's propagation direction is perpendicular to the substrate.)

The electromagnetic field representation for diffraction order $(m_1^{[\perp]},m_2^{[\perp]})$ will be based on $\hat{s}[m_1^{[\perp]},m_2^{[\perp]}]$ and $\hat{q}^{[\pm]}[m_1^{[\perp]},m_2^{[\perp]}]$ projections (unprimed) at a particular level $x_1=x_1^{[0]}$ in the grating, and will be based on $\hat{s}'[m_1^{[\perp]},m_2^{[\perp]}]$ and $\hat{q}'^{[\pm]}[m_1^{[\perp]},m_2^{[\perp]}]$ projections (primed) at level $x_1=x_1^{[1]}$ infinitesimally above $x_1^{[0]}$. ($x_1^{[0]}$ and $x_1^{[1]}$ are defined to fit the basis change into the framework of an S matrix, as described in section 7.) One of the two basis vectors $\hat{s}$ or $\hat{s}'$ is defined by the second branch condition ("if … = 0") in equation 6.39 (with $(m_1,m_2)=(m_1^{[\perp]},m_2^{[\perp]})$); and the other is similarly defined, but with $\vec{f}_1^{[s]}[l_1]$ substituted for $\vec{f}_1^{[g]}$ (as in equation 15.1). $\hat{q}^{[\pm]}$ and $\hat{q}'^{[\pm]}$ are defined according to equation 6.41.

The electromagnetic field is furthermore represented in terms of up and down waves in a fictitious medium of permittivity $\varepsilon^{[f]}$. Based on equations 6.50-53 (with $\varepsilon^{[c]}=\varepsilon^{[f]}$), the $\hat{s}$ basis change is described by the following relations in which "…" is an abbreviation for "$m_1^{[\perp]},m_2^{[\perp]}$",

$$\begin{aligned}
&\hat{s}[\ldots](\mathit{ff}E_s^{[+]}[\ldots][x_1^{[0]}]+\mathit{ff}E_s^{[-]}[\ldots][x_1^{[0]}])+\\
&\quad\hat{q}^{[+]}[\ldots]\frac{\lambda}{\varepsilon^{[f]}}f_1^{[+]}[\ldots](\mathit{ff}H_s^{[+]}[\ldots][x_1^{[0]}]+\mathit{ff}H_s^{[-]}[\ldots][x_1^{[0]}])=\\
&\hat{s}'[\ldots](\mathit{ff}E_s^{[+]}[\ldots][x_1^{[1]}]+\mathit{ff}E_s^{[-]}[\ldots][x_1^{[1]}])+\\
&\quad\hat{q}'^{[+]}[\ldots]\frac{\lambda}{\varepsilon^{[f]}}f_1^{[+]}[\ldots](\mathit{ff}H_s^{[+]}[\ldots][x_1^{[1]}]+\mathit{ff}H_s^{[-]}[\ldots][x_1^{[1]}])
\end{aligned} \tag{15.3}$$

$$\begin{aligned}
&\hat{s}[\ldots](\mathit{ff}H_s^{[+]}[\ldots][x_1^{[0]}]-\mathit{ff}H_s^{[-]}[\ldots][x_1^{[0]}])-\\
&\quad\hat{q}^{[+]}[\ldots]\lambda\, f_1^{[+]}[\ldots](\mathit{ff}E_s^{[+]}[\ldots][x_1^{[0]}]-\mathit{ff}E_s^{[-]}[\ldots][x_1^{[0]}])=\\
&\hat{s}'[\ldots](\mathit{ff}H_s^{[+]}[\ldots][x_1^{[1]}]-\mathit{ff}H_s^{[-]}[\ldots][x_1^{[1]}])-\\
&\quad\hat{q}'^{[+]}[\ldots]\lambda\, f_1^{[+]}[\ldots](\mathit{ff}E_s^{[+]}[\ldots][x_1^{[1]}]-\mathit{ff}E_s^{[-]}[\ldots][x_1^{[1]}])
\end{aligned} \tag{15.4}$$

(These equations assert continuity of the grating-tangential projections of $\mathit{ff}\vec{E}[m_1^{[\perp]},m_2^{[\perp]}][x_1]$ and $\mathit{ff}\vec{H}[m_1^{[\perp]},m_2^{[\perp]}][x_1]$ between $x_1^{[0]}$ and $x_1^{[1]}$.) The $f_1^{[+]}$ term is defined by equations 6.35 and 15.2,

$$f_1^{[+]}[\ldots] = \frac{\sqrt{\varepsilon^{[f]}}}{\lambda} \tag{15.5}$$

and the $\hat{s}$ and $\hat{q}^{[+]}$ terms have the coordinate projections defined by equations 6.38 and 6.41. Making these substitutions, equations 15.3 and 15.4 can be restated as

$$\begin{pmatrix} s_2[\ldots] & s_3[\ldots]/\sqrt{\varepsilon^{[f]}} & -s_2'[\ldots] & -s_3'[\ldots]/\sqrt{\varepsilon^{[f]}} \\ s_3[\ldots] & -s_2[\ldots]/\sqrt{\varepsilon^{[f]}} & -s_3'[\ldots] & s_2'[\ldots]/\sqrt{\varepsilon^{[f]}} \\ s_3[\ldots]\sqrt{\varepsilon^{[f]}} & -s_2[\ldots] & s_3'[\ldots]\sqrt{\varepsilon^{[f]}} & -s_2'[\ldots] \\ -s_2[\ldots]\sqrt{\varepsilon^{[f]}} & -s_3[\ldots] & -s_2'[\ldots]\sqrt{\varepsilon^{[f]}} & -s_3'[\ldots] \end{pmatrix} \begin{pmatrix} ffE_s^{[-]}[\ldots][x_1^{[0]}] \\ ffH_s^{[-]}[\ldots][x_1^{[0]}] \\ ffE_s^{[+]}[\ldots][x_1^{[1]}] \\ ffH_s^{[+]}[\ldots][x_1^{[1]}] \end{pmatrix} =$$
$$\begin{pmatrix} -s_2[\ldots] & -s_3[\ldots]/\sqrt{\varepsilon^{[f]}} & s_2'[\ldots] & s_3'[\ldots]/\sqrt{\varepsilon^{[f]}} \\ -s_3[\ldots] & s_2[\ldots]/\sqrt{\varepsilon^{[f]}} & s_3'[\ldots] & -s_2'[\ldots]/\sqrt{\varepsilon^{[f]}} \\ s_3[\ldots]\sqrt{\varepsilon^{[f]}} & -s_2[\ldots] & s_3'[\ldots]\sqrt{\varepsilon^{[f]}} & -s_2'[\ldots] \\ -s_2[\ldots]\sqrt{\varepsilon^{[f]}} & -s_3[\ldots] & -s_2'[\ldots]\sqrt{\varepsilon^{[f]}} & -s_3'[\ldots] \end{pmatrix} \begin{pmatrix} ffE_s^{[+]}[\ldots][x_1^{[0]}] \\ ffH_s^{[+]}[\ldots][x_1^{[0]}] \\ ffE_s^{[-]}[\ldots][x_1^{[1]}] \\ ffH_s^{[-]}[\ldots][x_1^{[1]}] \end{pmatrix} \tag{15.6}$$

If the j-th index block comprises diffraction order $(m_1^{[\perp]}, m_2^{[\perp]})$, then equation 15.6 can be converted to equation 9.3, wherein

$$F^{[\pm,j]}[x_1] = \begin{pmatrix} ffE_s^{[\pm]}[\ldots][x_1] \\ ffH_s^{[\pm]}[\ldots][x_1] \end{pmatrix} \tag{15.7}$$

$$S00^{[j]} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{15.8}$$

$$S01^{[j]} = \begin{pmatrix} (s_2[\ldots]s_2'[\ldots] + s_3[\ldots]s_3'[\ldots]) & (s_2[\ldots]s_3'[\ldots] - s_3[\ldots]s_2'[\ldots])/\sqrt{\varepsilon^{[f]}} \\ -(s_2[\ldots]s_3'[\ldots] - s_3[\ldots]s_2'[\ldots])\sqrt{\varepsilon^{[f]}} & (s_2[\ldots]s_2'[\ldots] + s_3[\ldots]s_3'[\ldots]) \end{pmatrix} \tag{15.9}$$

$$S10^{[j]} = \begin{pmatrix} (s_2[\ldots]s_2'[\ldots] + s_3[\ldots]s_3'[\ldots]) & -(s_2[\ldots]s_3'[\ldots] - s_3[\ldots]s_2'[\ldots])/\sqrt{\varepsilon^{[f]}} \\ (s_2[\ldots]s_3'[\ldots] - s_3[\ldots]s_2'[\ldots])\sqrt{\varepsilon^{[f]}} & (s_2[\ldots]s_2'[\ldots] + s_3[\ldots]s_3'[\ldots]) \end{pmatrix} \tag{15.10}$$

$$S11^{[j]} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{15.11}$$

The inverse S matrix transformation has the same form as equations 15.8-11, but with $\hat{s}[\ldots]$ and $\hat{s}'[\ldots]$ interchanged. Swapping $S01^{[j]}$ and $S10^{[j]}$ is equivalent to inverting the S matrix transformation. In implementing the grating stacking operations, the S matrix defined above must be applied to convert from the $\hat{s}$ basis defined by equation 13.90 to that of equation 15.1; then the uniperiodic stratum's S matrix (with TE and TM decoupled) is applied; and then the inverse of the above S matrix is applied to revert back to the $\hat{s}$ basis of equation 13.90.

16. The S matrix for a homogeneous stratum (with surfaces)

A homogeneous stratum is a special case of a periodic stratum in which the ranges of n_1 and n_2 in equation 13.1 are both limited to $\{0\}$. Denoting the stratum permittivity as $\varepsilon^{[c]}$, equation 13.1 reduces to

$$\varepsilon[\vec{x}] = \varepsilon 1[l_1][x_2, x_3] = f\!f\varepsilon 1[l_1, 0, 0] = \varepsilon^{[c]} \tag{16.1}$$

In this case, there is no diffractive coupling between the field's Fourier orders, so each decoupled index block is associated with a single Fourier order. Furthermore, there is no polarization coupling, so the enumeration indices k in section 9 and section 13 are limited to the range $\{1\}$; i.e.

$$k^{[\max]}[j] = 1 \tag{16.2}$$

(cf. equations 9.1 and 13.56). The polarization indices $P^{[\text{enum}]}[j,1]$ and wave amplitudes $F_1^{[\pm,j]}$ (equation 9.2) are defined as

$$P^{[\text{enum}]}[j,1] = \begin{cases} 0 & \text{for TE} \\ 1 & \text{for TM} \end{cases} \tag{16.3}$$

$$F_1^{[\pm,j]} = \begin{cases} f\!f E_{s,1}^{[\pm,j]} & \text{for TE} \\ f\!f H_{s,1}^{[\pm,j]} & \text{for TM} \end{cases} \tag{16.4}$$

The matrices $rtt\varepsilon 1$, $rtrt\varepsilon 1$, and $trtr\varepsilon 1$ that appear in equations 13.108 and 13.109 (and which are defined in equations 13.63, 13.82, and 13.83) simplify to 1-by-1 matrices,

$$rtt\varepsilon 1_{1,1} = 1/\varepsilon^{[c]} \tag{16.5}$$

$$rtrt\varepsilon 1_{1,1} = \varepsilon^{[c]} \tag{16.6}$$

$$trtr\varepsilon 1_{1,1} = \varepsilon^{[c]} \tag{16.7}$$

Differential equations 13.106 and 13.107 separate into decoupled TE and TM fields,

$$\begin{cases} \partial_1\, ffE_s^{[j]} = i\, DEH^{[\mathrm{TE}]}\, ffH_q^{[j]}, \quad \partial_1\, ffH_q^{[j]} = i\, DHE^{[\mathrm{TE}]}\, ffE_s^{[j]} & (\mathrm{TE}) \\ \partial_1\, ffE_q^{[j]} = i\, DEH^{[\mathrm{TM}]}\, ffH_s^{[j]}, \quad \partial_1\, ffH_s^{[j]} = i\, DHE^{[\mathrm{TM}]}\, ffE_q^{[j]} & (\mathrm{TM}) \end{cases} \tag{16.8}$$

wherein

$$DEH^{[TE]} = -\frac{2\pi}{\lambda} \tag{16.9}$$

$$DHE^{[TE]} = 2\pi\left(\lambda\left((df_2^{[p,j]})^2 + (df_3^{[p,j]})^2\right) - \frac{\varepsilon^{[c]}}{\lambda}\right) = -2\pi\,\lambda\,(f_1^{[+,j,c]})^2 \tag{16.10}$$

$$DEH^{[TM]} = 2\pi\left(\frac{1}{\lambda} - \frac{\lambda}{\varepsilon^{[c]}}\left((df_2^{[p,j]})^2 + (df_3^{[p,j]})^2\right)\right) = \frac{2\pi\,\lambda}{\varepsilon^{[c]}}(f_1^{[+,j,c]})^2 \tag{16.11}$$

$$DHE^{[TM]} = \frac{2\pi\,\varepsilon^{[c]}}{\lambda} \tag{16.12}$$

with $f_1^{[\pm,j,c]}$ defined as

$$f_1^{[\pm,j,c]} = \pm\sqrt{\frac{\varepsilon^{[c]}}{\lambda^2} - (f_{2,1}^{[p,j]})^2 - (f_{3,1}^{[p,j]})^2} \tag{16.13}$$

In this case, the matrices DEH and DHE (equations 13.108 and 13.109) are 2-by-2 matrices having the following form,

$$DEH = \begin{pmatrix} 0 & DEH^{[TE]} \\ DEH^{[TM]} & 0 \end{pmatrix} \tag{16.14}$$

$$DHE = \begin{pmatrix} 0 & DHE^{[TM]} \\ DHE^{[TE]} & 0 \end{pmatrix} \tag{16.15}$$

Solutions to equations 16.8 are of the form

$$ffE_s^{[j]}[x_1^{[0]} + \Delta x_1] = \Phi EE^{[\mathrm{TE}]}\; ffE_s^{[j]}[x_1^{[0]}] + i\,\Phi EH^{[\mathrm{TE}]}\; ffH_q^{[j]}[x_1^{[0]}] \tag{16.16}$$

$$ffH_q^{[j]}[x_1^{[0]} + \Delta x_1] = -i\,\Phi HE^{[\mathrm{TE}]}\; ffE_s^{[j]}[x_1^{[0]}] + \Phi HH^{[\mathrm{TE}]}\; ffH_q^{[j]}[x_1^{[0]}] \tag{16.17}$$

$$ffE_q^{[j]}[x_1^{[0]} + \Delta x_1] = \Phi EE^{[\mathrm{TM}]}\; ffE_q^{[j]}[x_1^{[0]}] + i\,\Phi EH^{[\mathrm{TM}]}\; ffH_s^{[j]}[x_1^{[0]}] \tag{16.18}$$

$$ffH_s^{[j]}[x_1^{[0]} + \Delta x_1] = -i\,\Phi HE^{[\mathrm{TM}]}\; ffE_q^{[j]}[x_1^{[0]}] + \Phi HH^{[\mathrm{TM}]}\; ffH_s^{[j]}[x_1^{[0]}] \tag{16.19}$$

wherein

$$\begin{pmatrix} \Phi EE^{[\mathrm{TE}]} & \Phi EH^{[\mathrm{TE}]} \\ \Phi HE^{[\mathrm{TE}]} & \Phi HH^{[\mathrm{TE}]} \end{pmatrix} = \exp\left[\begin{pmatrix} \boldsymbol{0} & DEH^{[\mathrm{TE}]} \\ -DHE^{[\mathrm{TE}]} & \boldsymbol{0} \end{pmatrix} \Delta x_1\right] \tag{16.20}$$

$$\begin{pmatrix} \Phi EE^{[\mathrm{TM}]} & \Phi EH^{[\mathrm{TM}]} \\ \Phi HE^{[\mathrm{TM}]} & \Phi HH^{[\mathrm{TM}]} \end{pmatrix} = \exp\left[\begin{pmatrix} \boldsymbol{0} & DEH^{[\mathrm{TM}]} \\ -DHE^{[\mathrm{TM}]} & \boldsymbol{0} \end{pmatrix} \Delta x_1\right] \tag{16.21}$$

The exponential matrices can be evaluated in closed form,

$$\begin{pmatrix} \Phi EE^{[\mathrm{TE}]} & \Phi EH^{[\mathrm{TE}]} \\ \Phi HE^{[\mathrm{TE}]} & \Phi HH^{[\mathrm{TE}]} \end{pmatrix} = \begin{pmatrix} \cos[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] & -\dfrac{1}{\lambda\, f_1^{[+,j,c]}} \sin[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] \\ \lambda\, f_1^{[+,j,c]} \sin[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] & \cos[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] \end{pmatrix} \tag{16.22}$$

$$\begin{pmatrix} \Phi EE^{[\mathrm{TM}]} & \Phi EH^{[\mathrm{TM}]} \\ \Phi HE^{[\mathrm{TM}]} & \Phi HH^{[\mathrm{TM}]} \end{pmatrix} = \begin{pmatrix} \cos[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] & \dfrac{\lambda\, f_1^{[+,j,c]}}{\varepsilon^{[c]}} \sin[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] \\ -\dfrac{\varepsilon^{[c]}}{\lambda\, f_1^{[+,j,c]}} \sin[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] & \cos[2\pi\, f_1^{[+,j,c]}\, \Delta x_1] \end{pmatrix} \tag{16.23}$$

Equations 16.16-19 translate to an equation having the same form as equation 13.116, wherein

$$
\Phi^{[\pm\pm']} = \begin{cases} \dfrac{1}{2}\begin{pmatrix} \Phi EE^{[\mathrm{TE}]} \mp' i\,\Phi EH^{[\mathrm{TE}]}\,\lambda\, df_1^{[+,j]} \\ \pm(\lambda\, df_1^{[+,j]})^{-1}\left(i\,\Phi HE^{[\mathrm{TE}]} \pm' \Phi HH^{[\mathrm{TE}]}\,\lambda\, df_1^{[+,j]}\right) \end{pmatrix} & \text{for TE} \\ \dfrac{1}{2}\begin{pmatrix} \varepsilon^{[f]}\,(\lambda\, df_1^{[+,j]})^{-1}\left(\Phi EE^{[\mathrm{TM}]}\,(\lambda/\varepsilon^{[f]})\, df_1^{[+,j]} \pm' i\,\Phi EH^{[\mathrm{TM}]}\right) \\ \pm\left(-i\,\Phi HE^{[\mathrm{TM}]}\,(\lambda/\varepsilon^{[f]})\, df_1^{[+,j]} \pm' \Phi HH^{[\mathrm{TM}]}\right) \end{pmatrix} & \text{for TM} \end{cases}
\tag{16.24}
$$

Substituting from equation 16.22 and 16.23, equation 16.24 reduces to the following,

for TE :

$$
\Phi^{[\pm\pm']} = \frac{1}{2}\begin{pmatrix} \cos[2\pi\, f_1^{[+,j,c]}\,\Delta x_1] \pm' i\,\dfrac{f_1^{[+,j,f]}}{f_1^{[+,j,c]}}\sin[2\pi\, f_1^{[+,j,c]}\,\Delta x_1] \\ \pm\left(i\,\dfrac{f_1^{[+,j,c]}}{f_1^{[+,j,f]}}\sin[2\pi\, f_1^{[+,j,c]}\,\Delta x_1] \pm' \cos[2\pi\, f_1^{[+,j,c]}\,\Delta x_1]\right) \end{pmatrix}
\tag{16.25}
$$

for TM :

$$
\Phi^{[\pm\pm']} = \frac{1}{2}\begin{pmatrix} \cos[2\pi\, f_1^{[+,j,c]}\,\Delta x_1] \pm' i\,\dfrac{\varepsilon^{[f]}\, f_1^{[+,j,c]}}{\varepsilon^{[c]}\, f_1^{[+,j,f]}}\sin[2\pi\, f_1^{[+,j,c]}\,\Delta x_1] \\ \pm\left(i\,\dfrac{\varepsilon^{[c]}\, f_1^{[+,j,f]}}{\varepsilon^{[f]}\, f_1^{[+,j,c]}}\sin[2\pi\, f_1^{[+,j,c]}\,\Delta x_1] \pm' \cos[2\pi\, f_1^{[+,j,c]}\,\Delta x_1]\right) \end{pmatrix}
\tag{16.26}
$$

The S matrix is then defined as in equation 13.120.

Note that $f_1^{[+,j,c]}$ can be zero in equations 16.25 and 16.26, but the ratio $\sin[2\pi\, f_1^{[+,j,c]}\,\Delta x_1]/ f_1^{[+,j,c]}$ reduces to $2\pi\,\Delta x_1$ when $f_1^{[+,j,c]}$ approaches zero[5]. If $f_1^{[+,j,f]}$ is zero the equations contain infinities, but this possibility is precluded by defining $\varepsilon^{[f]}$ to have a non-zero imaginary part.

[5] For numerical applications, the ratio $\sin[2\pi\, f_1^{[+,j,c]}\,\Delta x_1]/ f_1^{[+,j,c]}$ should be replaced by $2\pi\,\Delta x_1$ when $\left|2\pi\, f_1^{[+,j,c]}\,\Delta x_1\right|^2 < 6\,\delta$, wherein δ is the numeric precision (i.e., the smallest value such that 1 and $1+\delta$ are numerically distinguishable).

17. The reflection and transmission matrices

The j, k indices in equation 9.1 corresponding to the incident E field will be denoted as jiE, kiE; and the indices corresponding to the incident H field will be similarly denoted as jiH, kiH:

$$(m_1^{[\text{enum}]}[jiE,kiE], m_2^{[\text{enum}]}[jiE,kiE], P^{[\text{enum}]}[jiE,kiE]) = (0,0,0) \tag{17.1}$$

$$(m_1^{[\text{enum}]}[jiH,kiH], m_2^{[\text{enum}]}[jiH,kiH], P^{[\text{enum}]}[jiH,kiH]) = (0,0,1) \tag{17.2}$$

Taking the S matrix in equation 9.4 to represent the cumulative S matrix for the entire grating stack, the only non-zero terms on the right side of the equation are those corresponding to the incident field,

$$\begin{pmatrix} F_k^{[-,j]}[x_1^{[0]}] \\ F_k^{[+,j]}[x_1^{[1]}] \end{pmatrix} = \left(\begin{cases} \begin{pmatrix} S01_{k,kiE}^{[j]} \\ S11_{k,kiE}^{[j]} \end{pmatrix} F_{kiE}^{[-,j]}[x_1^{[1]}] & \text{if } j = jiE \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiE \end{cases} \right) + \left(\begin{cases} \begin{pmatrix} S01_{k,kiH}^{[j]} \\ S11_{k,kiH}^{[j]} \end{pmatrix} F_{kiH}^{[-,j]}[x_1^{[1]}] & \text{if } j = jiH \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiH \end{cases} \right) \tag{17.3}$$

The $x_1^{[0]}$ and $x_1^{[1]}$ values in this equation represent the x_1 coordinate below and above the grating, respectively,

$$x_1^{[0]} = b_1[0] - 0 \tag{17.4}$$

$$x_1^{[1]} = b_1[L_1] + 0 \tag{17.5}$$

(cf. equations 3.2 and 3.3).

Applying equation 9.2, equation 17.3 reduces to the following two equations, If $P^{[\mathrm{enum}]}[j,k]=0$,

$$\begin{pmatrix} ffE_s^{[-]}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]][x_1^{[0]}] \\ ffE_s^{[+]}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]][x_1^{[1]}] \end{pmatrix} = \left(\begin{cases} \begin{pmatrix} S01_{k,kiE}^{[j]} \\ S11_{k,kiE}^{[j]} \end{pmatrix} ffE_s^{[-]}[0,0][x_1^{[1]}] & \text{if } j = jiE \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiE \end{cases} \right) + \left(\begin{cases} \begin{pmatrix} S01_{k,kiH}^{[j]} \\ S11_{k,kiH}^{[j]} \end{pmatrix} ffH_s^{[-]}[0,0][x_1^{[1]}] & \text{if } j = jiH \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiH \end{cases} \right) \tag{17.6}$$

If $P^{[\mathrm{enum}]}[j,k]=1$,

$$\begin{pmatrix} ffH_s^{[-]}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]][x_1^{[0]}] \\ ffH_s^{[+]}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]][x_1^{[1]}] \end{pmatrix} = \left(\begin{cases} \begin{pmatrix} S01_{k,kiE}^{[j]} \\ S11_{k,kiE}^{[j]} \end{pmatrix} ffE_s^{[-]}[0,0][x_1^{[1]}] & \text{if } j = jiE \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiE \end{cases} \right) + \left(\begin{cases} \begin{pmatrix} S01_{k,kiH}^{[j]} \\ S11_{k,kiH}^{[j]} \end{pmatrix} ffH_s^{[-]}[0,0][x_1^{[1]}] & \text{if } j = jiH \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiH \end{cases} \right) \tag{17.7}$$

The H-field amplitudes $ffH_s^{[\pm]}$ are related to the E-field amplitudes $ffE_p^{[\pm]}$ by equation 6.48,

$$ffH_s^{[-]}[0,0][x_1^{[1]}] = \sqrt{\varepsilon^{[\mathrm{sup}]}}\; ffE_p^{[-]}[0,0][x_1^{[1]}] \tag{17.8}$$

$$ffH_s^{[-]}[m_1,m_2][x_1^{[0]}] = \sqrt{\varepsilon^{[\mathrm{sub}]}}\; ffE_p^{[-]}[m_1,m_2][x_1^{[0]}] \tag{17.9}$$

$$ffH_s^{[+]}[m_1,m_2][x_1^{[1]}] = \sqrt{\varepsilon^{[\mathrm{sup}]}}\; ffE_p^{[+]}[m_1,m_2][x_1^{[1]}] \tag{17.10}$$

wherein $\varepsilon^{[\mathrm{sub}]}$ and $\varepsilon^{[\mathrm{sup}]}$ are the substrate and superstrate permittivities (equations 3.11, 3.12). The field quantities in the above equations are designated as follows in equations 4.33 and 4.34,

$$E_s^{[i]}[\hat{e}_1\, x_1^{[1]}] = \mathit{ff}E_s^{[-]}[0,0][x_1^{[1]}] \tag{17.11}$$

$$E_p^{[i]}[\hat{e}_1\, x_1^{[1]}] = \mathit{ff}E_p^{[-]}[0,0][x_1^{[1]}] \tag{17.12}$$

$$\mathit{ff}E_s^{[r]}[m_1,m_2][\hat{e}_1\, x_1^{[1]}] = \mathit{ff}E_s^{[+]}[m_1,m_2][x_1^{[1]}] \tag{17.13}$$

$$\mathit{ff}E_p^{[r]}[m_1,m_2][\hat{e}_1\, x_1^{[1]}] = \mathit{ff}E_p^{[+]}[m_1,m_2][x_1^{[1]}] \tag{17.14}$$

$$\mathit{ff}E_s^{[t]}[m_1,m_2][\hat{e}_1\, x_1^{[0]}] = \mathit{ff}E_s^{[-]}[m_1,m_2][x_1^{[0]}] \tag{17.15}$$

$$\mathit{ff}E_p^{[t]}[m_1,m_2][\hat{e}_1\, x_1^{[0]}] = \mathit{ff}E_p^{[-]}[m_1,m_2][x_1^{[0]}] \tag{17.16}$$

Making these substitutions in equations 17.6 and 17.7, and comparing with equations 4.33 and 4.34, we obtain the following,

$$\begin{aligned}
&\text{If } P^{[\mathrm{enum}]}[j,k]=0,\\
&\begin{pmatrix} \mathit{ff}E_s^{[t]}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]][\hat{e}_1\, x_1^{[0]}] \\ \mathit{ff}E_s^{[r]}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]][\hat{e}_1\, x_1^{[1]}] \end{pmatrix} = \\
&\quad \left(\begin{cases} \begin{pmatrix} S01_{k,kiE}^{[j]} \\ S11_{k,kiE}^{[j]} \end{pmatrix} E_s^{[i]}[\hat{e}_1\, x_1^{[1]}] & \text{if } j = jiE \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiE \end{cases} \right) + \\
&\quad \left(\begin{cases} \begin{pmatrix} \sqrt{\varepsilon^{[\mathrm{sup}]}}\; S01_{k,kiH}^{[jiE]} \\ \sqrt{\varepsilon^{[\mathrm{sup}]}}\; S11_{k,kiH}^{[jiE]} \end{pmatrix} E_p^{[i]}[\hat{e}_1\, x_1^{[1]}] & \text{if } j = jiH \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiH \end{cases} \right) \\
&= \begin{pmatrix} T_{s,s}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] & T_{s,p}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] \\ R_{s,s}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] & R_{s,p}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] \end{pmatrix} \begin{pmatrix} E_s^{[i]}[\hat{e}_1\, x_1^{[1]}] \\ E_p^{[i]}[\hat{e}_1\, x_1^{[1]}] \end{pmatrix}
\end{aligned} \tag{17.17}$$

If $P^{[\text{enum}]}[j,k]=1$,

$$\begin{pmatrix} ffE_p^{[t]}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]][\hat{e}_1\, x_1^{[0]}] \\ ffE_p^{[r]}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]][\hat{e}_1\, x_1^{[1]}] \end{pmatrix} =$$

$$\left(\begin{cases} \begin{pmatrix} (1/\sqrt{\varepsilon^{[\text{sub}]}})\, S01_{k,kiE}^{[jiH]} \\ (1/\sqrt{\varepsilon^{[\text{sup}]}})\, S11_{k,kiE}^{[jiH]} \end{pmatrix} E_s^{[i]}[\hat{e}_1\, x_1^{[1]}] & \text{if } j = jiE \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiE \end{cases} \right) +$$

$$\left(\begin{cases} \begin{pmatrix} \sqrt{\varepsilon^{[\text{sup}]}/\varepsilon^{[\text{sub}]}}\, S01_{k,kiH}^{[jiH]} \\ S11_{k,kiH}^{[jiH]} \end{pmatrix} E_p^{[i]}[\hat{e}_1\, x_1^{[1]}] & \text{if } j = jiH \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiH \end{cases} \right)$$

$$= \begin{pmatrix} T_{p,s}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] & T_{p,p}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] \\ R_{p,s}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] & R_{p,p}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] \end{pmatrix} \begin{pmatrix} E_s^{[i]}[\hat{e}_1\, x_1^{[1]}] \\ E_p^{[i]}[\hat{e}_1\, x_1^{[1]}] \end{pmatrix} \tag{17.18}$$

Hence, the reflection and transmission matrices are defined as follows,

If $P^{[\text{enum}]}[j,k]=0$,

$$\begin{pmatrix} T_{s,s}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] \\ R_{s,s}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] \end{pmatrix} = \begin{cases} \begin{pmatrix} S01_{k,kiE}^{[j]} \\ S11_{k,kiE}^{[j]} \end{pmatrix} & \text{if } j = jiE \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiE \end{cases}$$

$$\begin{pmatrix} T_{s,p}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] \\ R_{s,p}[m_1^{[\text{enum}]}[j,k],m_2^{[\text{enum}]}[j,k]] \end{pmatrix} = \begin{cases} \begin{pmatrix} \sqrt{\varepsilon^{[\text{sup}]}}\, S01_{k,kiH}^{[jiE]} \\ \sqrt{\varepsilon^{[\text{sup}]}}\, S11_{k,kiH}^{[jiE]} \end{pmatrix} & \text{if } j = jiH \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j \neq jiH \end{cases} \tag{17.19}$$

If $P^{[\mathrm{enum}]}[j,k]=1$,

$$\begin{pmatrix} T_{p,s}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] \\ R_{p,s}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] \end{pmatrix} = \begin{cases} \begin{pmatrix} (1/\sqrt{\varepsilon^{[\mathrm{sub}]}})\, S01_{k,kiE}^{[jiH]} \\ (1/\sqrt{\varepsilon^{[\mathrm{sup}]}})\, S11_{k,kiE}^{[jiH]} \end{pmatrix} & \text{if } j=jiE \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j\neq jiE \end{cases}$$

$$\begin{pmatrix} T_{p,p}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] \\ R_{p,p}[m_1^{[\mathrm{enum}]}[j,k],m_2^{[\mathrm{enum}]}[j,k]] \end{pmatrix} = \begin{cases} \begin{pmatrix} \sqrt{\varepsilon^{[\mathrm{sup}]}/\varepsilon^{[\mathrm{sub}]}}\, S01_{k,kiH}^{[jiH]} \\ S11_{k,kiH}^{[jiH]} \end{pmatrix} & \text{if } j=jiH \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } j\neq jiH \end{cases} \tag{17.20}$$

Appendix A. Fourier expansion of the permittivity

The following description of the permittivity distribution in stratum l_1 assumes a coordinate orientation with $\hat{e}_2$ parallel to $\vec{f}_1^{[s]}[l_1]$ (equations 13.14). The stratum stripes for a biperiodic stratum are parallel to the period vector $\vec{d}_2^{[s]}[l_1]$ (e.g., see Figure 3), and $\vec{d}_2^{[s]}[l_1]$ is orthogonal to $\vec{f}_1^{[s]}[l_1]$ (equation 3.30); so the stripes are parallel to $\hat{e}_3$,

$$d_{2,2}^{[s]}[l_1]=0; \quad \vec{d}_2^{[s]}[l_1]=\hat{e}_3\, d_{3,2}^{[s]}[l_1] \quad \text{(biperiodic)} \tag{A.1}$$

Figure 10 illustrates the relationship between the stratum's period vectors ($\vec{d}_1^{[s]}[l_1]$, $\vec{d}_2^{[s]}[l_1]$), its basis frequency vectors ($\vec{f}_1^{[s]}[l_1]$, $\vec{f}_2^{[s]}[l_1]$), and the coordinate bases $\hat{e}_2$ and $\hat{e}_3$.

For a uniperiodic stratum $\vec{f}_1^{[s]}[l_1]$ is parallel to $\vec{d}_1^{[s]}[l_1]$ (equation 3.23); hence equations 13.14 imply

$$d_{3,1}^{[s]}[l_1]=0; \quad \vec{d}_1^{[s]}[l_1]=\hat{e}_2\, d_{2,1}^{[s]}[l_1] \quad \text{(uniperiodic)} \tag{A.2}$$

In this case, $\vec{d}_1^{[s]}[l_1]$ is orthogonal to the stripe orientation; so the stripes are also parallel to $\hat{e}_3$ for the uniperiodic case.

Within stratum l_1, the grating permittivity is independent of x_1, and in the case of a homogeneous stratum is constant; hence the permittivity has the functional form

$$\varepsilon[\vec{x}] = \begin{cases} \varepsilon 1[l_1][x_2, x_3] & \text{(periodic stratum)} \\ \varepsilon 1[l_1] & \text{(homogeneous stratum)} \end{cases} \tag{A.3}$$
$$\text{for } b_1[l_1 - 1] < x_1 < b_1[l_1],\ l_1 = 1, \ldots L_1$$

wherein the stratum boundaries are at $x_1 = b_1[0], b_1[1], \ldots$; and L_1 is the number of strata (equations 3.1 and 3.10).

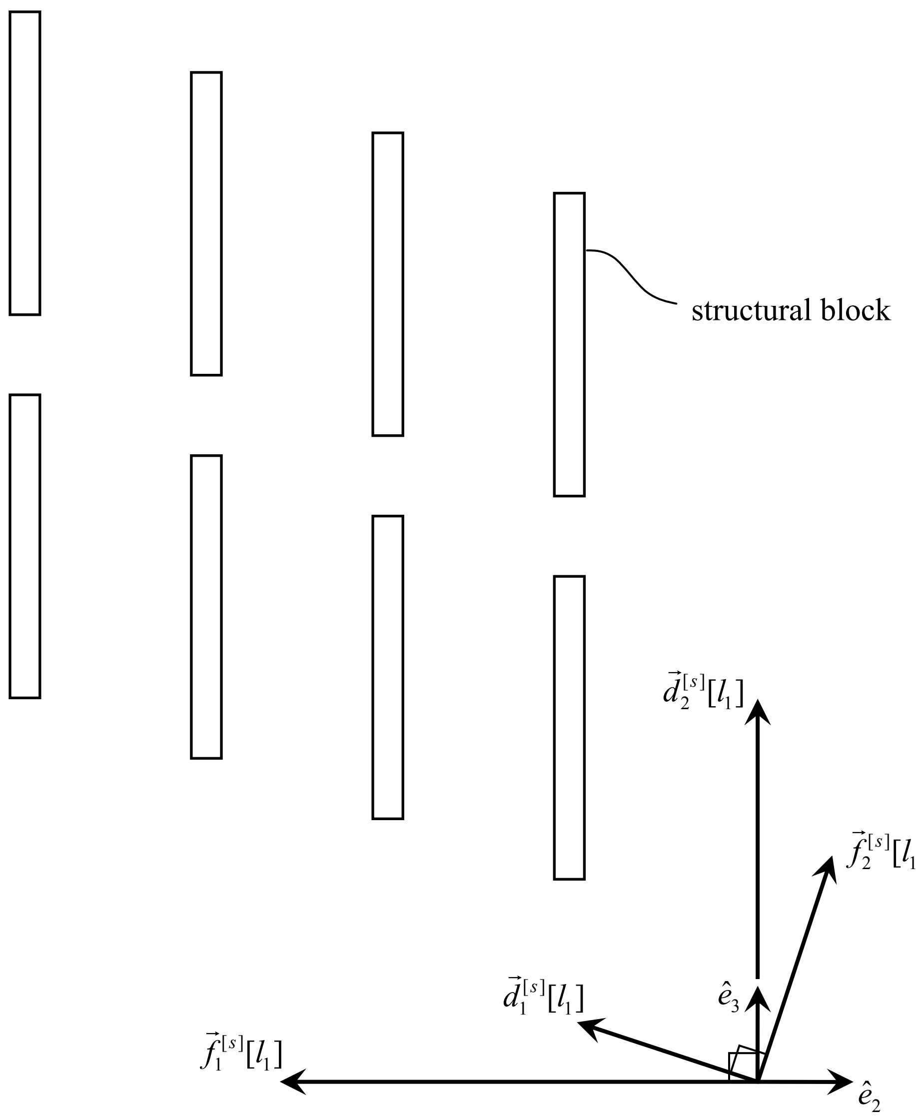

Figure 10. Stratum period vectors, basis frequency vectors, and coordinate bases.

A periodic stratum is partitioned into stripes. The permittivity within each stripe is independent of x_2, and in the case of a uniperiodic stratum is constant,

$$\varepsilon 1[l_1][x_2,x_3] = \begin{cases} \varepsilon 2[l_1,l_2][x_3] & \text{(biperiodic stratum)} \\ \varepsilon 2[l_1,l_2] & \text{(uniperiodic stratum)} \end{cases} \tag{A.4}$$
$$\quad \text{for } x_2 \text{ between } b_2[l_1,l_2-1] \text{ and } b_2[l_1,l_2]$$

wherein the stripe boundaries are at $x_2 = \ldots, b_2[l_1,0], b_2[l_1,1], \ldots$. (The boundary positions may be sorted in either increasing or decreasing order.)

Each stripe of a biperiodic stratum is either homogeneous, or is partitioned into homogeneous structural blocks,

$$\begin{aligned} &\varepsilon 2[l_1,l_2][x_3] = \varepsilon 3[l_1,l_2,l_3] \\ &\quad \text{for } x_3 \text{ between } b_3[l_1,l_2,l_3-1] \text{ and } b_3[l_1,l_2,l_3] \end{aligned} \tag{A.5}$$

wherein the block boundaries are at $x_3 = \ldots, b_3[l_1,l_2,0], b_3[l_1,l_2,1], \ldots$ (cf. equation 3.32).

The stripe boundaries $b_2[l_1,l_2]$ are defined in terms of the dimensionless parameters $c_1[l_1,l_2]$, which satisfy relations 3.33 and 3.34:

$$b_2[l_1,l_2] = c_1[l_1,l_2]\, d_{2,1}^{[s]}[l_1], \quad l_2 = 1,\ldots,L_2[l_1] \tag{A.6}$$

$$b_2[l_1,l_2+L_2[l_1]] = b_2[l_1,l_2] + d_{2,1}^{[s]}[l_1] \tag{A.7}$$

The block boundaries $b_3[l_1,l_2,l_3]$ within each inhomogeneous stripe are defined in terms of the dimensionless parameters $c_2[l_1,l_2,l_3]$, which satisfy relations 3.35-37:

$$\begin{aligned} &b_3[l_1,l_2,l_3] = c_1[l_1,l_2]\, d_{3,1}^{[s]}[l_1] + c_2[l_1,l_2,l_3]\, d_{3,2}^{[s]}[l_1], \\ &\quad l_2 = 1,\ldots,L_2[l_1], \quad l_3 = 1,\ldots,L_3[l_1,l_2] \end{aligned} \tag{A.8}$$

$$b_3[l_1,l_2+L_2[l_1],l_3] = b_3[l_1,l_2,l_3] + d_{3,1}^{[s]}[l_1] \tag{A.9}$$

$$b_3[l_1,l_2,l_3+L_3[l_1,l_2]] = b_3[l_1,l_2,l_3] + d_{3,2}^{[s]}[l_1] \tag{A.10}$$

A biperiodic stratum is periodic in x_3,

$$\varepsilon 1[l_1][x_2,x_3+d_{3,2}^{[s]}[l_1]] = \varepsilon 1[l_1][x_2,x_3] \tag{A.11}$$

(from equations 3.16 and A.1); hence the permittivity has a Fourier expansion of the form

$$\varepsilon 1[l_1][x_2, x_3] = \sum_{n_2} f\varepsilon 1[l_1, n_2][x_2] \exp[i\,2\pi\, n_2\, f_{3,2}^{[s]}[l_1]\, x_3] \tag{A.12}$$

wherein

$$f_{3,2}^{[s]}[l_1] = 1/d_{3,2}^{[s]}[l_1] \tag{A.13}$$

(cf. equation 13.15). Equation A.4 implies that $f\varepsilon 1[l_1, n_2][x_2]$ is of the form

$$\begin{aligned} &f\varepsilon 1[l_1, n_2][x_2] = f\varepsilon 2[l_1, l_2, n_2] \\ &\quad \text{for } x_2 \text{ between } b_2[l_1, l_2 - 1] \text{ and } b_2[l_1, l_2] \end{aligned} \tag{A.14}$$

wherein

$$\varepsilon 2[l_1, l_2][x_3] = \sum_{n_2} f\varepsilon 2[l_1, l_2, n_2] \exp[i\,2\pi\, n_2\, f_{3,2}^{[s]}[l_1]\, x_3] \tag{A.15}$$

Equation A.5 is substituted on the left side of equation A.15, and the Fourier inversion formula is applied to determine the Fourier coefficients $f\varepsilon 2[l_1, l_2, n_2]$,

$$\begin{aligned} f\varepsilon 2[l_1, l_2, n_2] &= f_{3,2}^{[s]}[l_1] \int_{b_3[l_1, l_2, 0]}^{b_3[l_1, l_2, L_3[l_1, l_2]]} \varepsilon 2[l_1, l_2][x_3] \exp[-i\,2\pi\, n_2\, f_{3,2}^{[s]}[l_1]\, x_3]\, dx_3 \\ &= \sum_{l_3=1}^{L_3[l_1, l_2]} \left(\begin{aligned} &\varepsilon 3[l_1, l_2, l_3]\, f_{3,2}^{[s]}[l_1] (b_3[l_1, l_2, l_3] - b_3[l_1, l_2, l_3 - 1]) \cdot \\ &\quad \operatorname{sinc}[\pi\, n_2\, f_{3,2}^{[s]}[l_1] (b_3[l_1, l_2, l_3] - b_3[l_1, l_2, l_3 - 1])] \cdot \\ &\quad \exp[-i\,\pi\, n_2\, f_{3,2}^{[s]}[l_1] (b_3[l_1, l_2, l_3] + b_3[l_1, l_2, l_3 - 1])] \end{aligned} \right) \end{aligned} \tag{A.16}$$

wherein the sinc function is defined as[6]

$$\operatorname{sinc}[x] = \begin{cases} \sin[x]/x & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases} \tag{A.17}$$

For a uniperiodic stratum, equation A.15 is simply

---

[6] For numerical applications the following implementation of the sinc function is preferred over definition A.17, in order to avoid excessive round-off errors for small x:

$$\operatorname{sinc}[x] = \begin{cases} \sin[x]/x & \text{if } x^2 \geq 6\,\delta \\ 1 & \text{if } x^2 < 6\,\delta \end{cases}$$

wherein δ is the numeric precision (i.e., the smallest value such that 1 and $1+\delta$ are numerically distinguishable).

$$\varepsilon 2[l_1,l_2][x_3] = f\varepsilon 2[l_1,l_2,0] \quad \text{(uniperiodic)} \tag{A.18}$$

Equation A.16 can be recast in a form that is independent of the $\hat{e}_2$, $\hat{e}_3$ coordinate orientation by using the following identity,

$$\begin{aligned} f_{3,2}^{[s]}[l_1]\, b_3[l_1,l_2,l_3] &= (\vec{f}_2^{[s]}[l_1] \bullet \hat{e}_3)((c_1[l_1,l_2]\, \vec{d}_1^{[s]}[l_1] + c_2[l_1,l_2,l_3]\, \vec{d}_2^{[s]}[l_1]) \bullet \hat{e}_3) \\ &= \frac{(\vec{f}_2^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1])((c_1[l_1,l_2]\, \vec{d}_1^{[s]}[l_1] + c_2[l_1,l_2,l_3]\, \vec{d}_2^{[s]}[l_1]) \bullet \vec{d}_2^{[s]}[l_1])}{\vec{d}_2^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1]} \\ &= c_1[l_1,l_2]\, \frac{\vec{d}_1^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1])}{\vec{d}_2^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1]} + c_2[l_1,l_2,l_3] \end{aligned} \tag{A.19}$$

(from equations A.8, A.1, and 3.30). Defining

$$\gamma[l_1] = \frac{\vec{d}_1^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1])}{\vec{d}_2^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1]} \tag{A.20}$$

equation A.16 simplifies to

$$\begin{aligned} & f\varepsilon 2[l_1,l_2,n_2] = \\ & \sum_{l_3=1}^{L_3[l_1,l_2]} \left(\begin{aligned} & \varepsilon 3[l_1,l_2,l_3](c_2[l_1,l_2,l_3] - c_2[l_1,l_2,l_3-1]) \cdot \\ & \quad \operatorname{sinc}[\pi\, n_2\, (c_2[l_1,l_2,l_3] - c_2[l_1,l_2,l_3-1])] \cdot \\ & \quad \exp[-i\,\pi\, n_2\, (c_2[l_1,l_2,l_3] + c_2[l_1,l_2,l_3-1] + 2\, c_1[l_1,l_2]\gamma[l_1])] \end{aligned} \right) \end{aligned} \tag{A.21}$$

(For a uniperiodic stratum $\vec{d}_2^{[s]}[l_1]$ is implicitly of infinite magnitude and perpendicular to $\vec{d}_1^{[s]}[l_1]$, so $\gamma[l_1]$ is zero.)

The periodicity condition 3.15, applied to the right side of equation A.12, yields the x_2-periodicity condition 13.43 for $f\varepsilon 1$,

$$\begin{aligned} & f\varepsilon 1[l_1,n_2][x_2 + 1/f_{2,1}^{[s]}[l_1]]\exp[-i\, 2\pi\, n_2\, f_{2,2}^{[s]}[l_1](x_2 + 1/f_{2,1}^{[s]}[l_1])] = \\ & \quad f\varepsilon 1[l_1,n_2][x_2]\exp[-i\, 2\pi\, n_2\, f_{2,2}^{[s]}[l_1]\, x_2] \end{aligned} \tag{A.22}$$

wherein

$$f_{2,1}^{[s]}[l_1] = 1/d_{2,1}^{[s]}[l_1] \tag{A.23}$$

$$f_{2,2}^{[s]}[l_1] = -d_{3,1}^{[s]}[l_1] / (d_{2,1}^{[s]}[l_1]\, d_{3,2}^{[s]}[l_1]) \tag{A.24}$$

(Figure 10 illustrates the geometric relationships defined by conditions 13.14, A.1, A.13, A.23, and A.24; cf. Figure 3 and equation 3.30.) The Fourier expansion 13.16 for $f\varepsilon 1$ is obtained from equation A.22,

$$f\varepsilon 1[l_1, n_2][x_2] = \sum_{n_1} ff\varepsilon 1[l_1, n_1, n_2] \exp[i\, 2\pi\, (n_1\, f_{2,1}^{[s]}[l_1] + n_2\, f_{2,2}^{[s]}[l_1])\, x_2\,] \tag{A.25}$$

Substituting from equation A.14, the Fourier coefficients $ff\varepsilon 1[l_1, n_1, n_2]$ are obtained,

$$\begin{aligned} ff\varepsilon 1[l_1, n_1, n_2] &= f_{2,1}^{[s]}[l_1] \int_{b_2[l_1,0]}^{b_2[l_1, L_2[l_1]]} f\varepsilon 1[l_1, n_2][x_2] \exp[-i\, 2\pi\, (n_1\, f_{2,1}^{[s]}[l_1] + n_2\, f_{2,2}^{[s]}[l_1])\, x_2\,]\, dx_2 \\ &= \sum_{l_2=1}^{L_2[l_1]} \left(\begin{array}{l} f\varepsilon 2[l_1, l_2, n_2]\, f_{2,1}^{[s]}[l_1]\, (b_2[l_1, l_2] - b_2[l_1, l_2 - 1]) \cdot \\ \quad \operatorname{sinc}[\pi\, (n_1\, f_{2,1}^{[s]}[l_1] + n_2\, f_{2,2}^{[s]}[l_1])\, (b_2[l_1, l_2] - b_2[l_1, l_2 - 1])] \cdot \\ \quad \exp[-i\, \pi\, (n_1\, f_{2,1}^{[s]}[l_1] + n_2\, f_{2,2}^{[s]}[l_1])\, (b_2[l_1, l_2] + b_2[l_1, l_2 - 1])] \end{array} \right) \end{aligned} \tag{A.26}$$

Equation A.26 can be recast in a form that is independent of the $\hat{e}_2$, $\hat{e}_3$ coordinate orientation by using the following identities,

$$\begin{aligned} f_{2,1}^{[s]}[l_1]\, b_2[l_1, l_2] &= (\vec{f}_1^{[s]}[l_1] \bullet \hat{e}_2)\, (c_1[l_1, l_2]\, \vec{d}_1^{[s]}[l_1] \bullet \hat{e}_2) \\ &= c_1[l_1, l_2]\, \vec{d}_1^{[s]}[l_1] \bullet \vec{f}_1^{[s]}[l_1] = c_1[l_1, l_2] \end{aligned} \tag{A.27}$$

$$\begin{aligned} f_{2,2}^{[s]}[l_1]\, b_2[l_1, l_2] &= (\vec{f}_2^{[s]}[l_1] \bullet \hat{e}_2)\, (c_1[l_1, l_2]\, \vec{d}_1^{[s]}[l_1] \bullet \hat{e}_2) \\ &= \frac{(\vec{f}_2^{[s]}[l_1] \bullet \vec{f}_1^{[s]}[l_1])\, (c_1[l_1, l_2]\, \vec{d}_1^{[s]}[l_1] \bullet \vec{f}_1^{[s]}[l_1])}{\vec{f}_1^{[s]}[l_1] \bullet \vec{f}_1^{[s]}[l_1]} \\ &= c_1[l_1, l_2]\, \frac{\vec{f}_2^{[s]}[l_1] \bullet \vec{f}_1^{[s]}[l_1]}{\vec{f}_1^{[s]}[l_1] \bullet \vec{f}_1^{[s]}[l_1]} = -c_1[l_1, l_2]\, \frac{\vec{d}_1^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1])}{\vec{d}_2^{[s]}[l_1] \bullet \vec{d}_2^{[s]}[l_1]} \end{aligned} \tag{A.28}$$

(from equations A.6, 13.14, and 3.30). With these substitutions, equation A.26 simplifies to

$$\begin{aligned} & ff\varepsilon 1[l_1, n_1, n_2] = \\ & \sum_{l_2=1}^{L_2[l_1]} \left(\begin{array}{l} f\varepsilon 2[l_1, l_2, n_2]\, (c_1[l_1, l_2] - c_1[l_1, l_2 - 1]) \cdot \\ \quad \operatorname{sinc}[\pi\, (n_1 - n_2\, \gamma[l_1])\, (c_1[l_1, l_2] - c_1[l_1, l_2 - 1])] \cdot \\ \quad \exp[-i\, \pi\, (n_1 - n_2\, \gamma[l_1])\, (c_1[l_1, l_2] + c_1[l_1, l_2 - 1])] \end{array} \right) \end{aligned} \tag{A.29}$$

The Toeplitz matrix $t\varepsilon 1$ (definition 13.35) and its inverse $rt\varepsilon 1$ (implicit definition 13.37) have the following form (from equation A.14),

$$\begin{aligned} & t\varepsilon 1[l_1, n_2, n_2'][x_2] = f\varepsilon 1[l_1, n_2 - n_2'][x_2] = t\varepsilon 2[l_1, l_2, n_2, n_2'] \\ & \quad \text{for } x_2 \text{ between } b_2[l_1, l_2 - 1] \text{ and } b_2[l_1, l_2] \end{aligned} \tag{A.30}$$

$$\begin{aligned} & rt\varepsilon 1[l_1, n_2, n_2'][x_2] = rt\varepsilon 2[l_1, l_2, n_2, n_2'] \\ & \quad \text{for } x_2 \text{ between } b_2[l_1, l_2 - 1] \text{ and } b_2[l_1, l_2] \end{aligned} \tag{A.31}$$

wherein

$$t\varepsilon 2[l_1, l_2, n_2, n_2'] = f\varepsilon 2[l_1, l_2, n_2 - n_2'] \tag{A.32}$$

$$\sum_{n_2'} (rt\varepsilon 2[l_1, l_2, n_2, n_2'])\, (t\varepsilon 2[l_1, l_2, n_2', n_2'']) = \begin{cases} 1, & n_2 = n_2'' \\ 0, & n_2 \ne n_2'' \end{cases} \tag{A.33}$$

(The indices n_2, n_2', and n_2'' in the above equations are limited to the set $\mathcal{N}_2^{[j]}$, in accordance with condition 13.54). $rt\varepsilon 1$ can be represented by Fourier expansion 13.46,

$$\begin{aligned} & rt\varepsilon 1[l_1, n_2, n_2'][x_2] = \\ & \quad \sum_{n_1} frt\varepsilon 1[l_1, n_1, n_2, n_2'] \exp[i\, 2\pi\, (n_1\, f_{2,1}^{[s]}[l_1] + (n_2 - n_2')\, f_{2,2}^{[s]}[l_1])\, x_2] \end{aligned} \tag{A.34}$$

wherein the Fourier coefficients are determined from equation A.31,

$$\begin{aligned} & frt\varepsilon 1[l_1, n_1, n_2, n_2'] \\ & \quad = f_{2,1}^{[s]}[l_1] \int_{b_2[l_1, 0]}^{b_2[l_1, L_2[l_1]]} rt\varepsilon 1[l_1, n_2, n_2'][x_2] \exp[-i\, 2\pi\, (n_1\, f_{2,1}^{[s]}[l_1] + (n_2 - n_2')\, f_{2,2}^{[s]}[l_1])\, x_2]\, dx_2 \\ & \quad = \sum_{l_2=1}^{L_2[l_1]} \left(\begin{aligned} & rt\varepsilon 2[l_1, l_2, n_2, n_2']\, f_{2,1}^{[s]}[l_1]\, (b_2[l_1, l_2] - b_2[l_1, l_2 - 1]) \cdot \\ & \quad \operatorname{sinc}[\pi\, (n_1\, f_{2,1}^{[s]}[l_1] + (n_2 - n_2')\, f_{2,2}^{[s]}[l_1])\, (b_2[l_1, l_2] - b_2[l_1, l_2 - 1])] \cdot \\ & \quad \exp[-i\, \pi\, (n_1\, f_{2,1}^{[s]}[l_1] + (n_2 - n_2')\, f_{2,2}^{[s]}[l_1])\, (b_2[l_1, l_2] + b_2[l_1, l_2 - 1])] \end{aligned} \right) \end{aligned} \tag{A.35}$$

With substitutions A.27 and A.28, this simplifies to

$$
\begin{aligned}
&frt\varepsilon 1[l_1, n_1, n_2, n_2'] = \\
&\quad = \sum_{l_2=1}^{L_2[l_1]} \begin{pmatrix} rt\varepsilon 2[l_1, l_2, n_2, n_2'] (c_1[l_1, l_2] - c_1[l_1, l_2 - 1]) \cdot \\ \operatorname{sinc}[\pi (n_1 - (n_2 - n_2') \gamma[l_1]) (c_1[l_1, l_2] - c_1[l_1, l_2 - 1])] \cdot \\ \exp[-i \pi (n_1 - (n_2 - n_2') \gamma[l_1]) (c_1[l_1, l_2] + c_1[l_1, l_2 - 1])] \end{pmatrix}
\end{aligned}
\tag{A.36}
$$

The reciprocal permittivity is denoted as "$r\varepsilon$",

$$
r\varepsilon[\vec{x}] = \frac{1}{\varepsilon[\vec{x}]} \tag{A.37}
$$

Most of the above equations involving ε (equations A.3-5, 11, 12, 14-16, 18, 21, 22, 30-36) also apply with the symbolic substitution of "$r\varepsilon$" for "ε" (cf. equations 13.28 and 13.47).

Appendix B. Exponential matrix computation

The exponential of a matrix A can be computed using a scale-and-square method [Ref. 6],

$$
\exp[A] = \exp[A / 2^{sp}]^{2^{sp}} \tag{B.1}
$$

wherein the right-hand exponential is evaluated using a Padé approximation and the power is evaluated by squaring the exponential sp times. But as indicated in section 13 (after equation 13.112), this "scale-and-square" algorithm will be replaced by a "scale-and-stack" method in which the exponential $\exp[A / 2^{sp}]$ is converted directly into an S matrix, which is then stacked with itself sp times to build up a stratum's full S matrix. (The stacking operations are more numerically stable than matrix squaring.)

In the context of equation 13.112, the matrix A is of the form

$$
A = \begin{pmatrix} \boldsymbol{0} & P \\ Q & \boldsymbol{0} \end{pmatrix} \tag{B.2}
$$

wherein

$$
P = DEH\,(x_1^{(1)} - x_1^{(0)}) \tag{B.3}
$$

$$
Q = -DHE\,(x_1^{(1)} - x_1^{(0)}) \tag{B.4}
$$

(A has a similar form in the context of equations 14.22 and 14.23.) The zero blocks in equation B.2 can be exploited to improve computational efficiency in two ways: First, the scaling power sp does not need to be as large as it would have to be without the zero blocks. And second, the number of operations in the Padé approximation can be reduced by skipping operations that multiply zero.

In accordance with Ref. 6, section 11, the scaling power sp can be chosen so that

$$\left\| A/2^{sp} \right\|_\infty < 1/2 \tag{B.5}$$

wherein $\left\| \ldots \right\|_\infty$ denotes the infinity-norm, i.e., the maximum absolute row sum,

$$\left\| A/2^{sp} \right\|_\infty = \max_m \sum_n | A_{m,n} | / 2^{sp} \tag{B.6}$$

It is evident from Eqations B.6 and B.2 that

$$\left\| A/2^{sp} \right\|_\infty = \max[\left\| P/2^{sp} \right\|_\infty, \left\| Q/2^{sp} \right\|_\infty] \tag{B.7}$$

However, a similarity transformation can be applied to the right-hand exponential in equation B.1 to obtain

$$\exp\left[\begin{pmatrix} \boldsymbol{0} & P \\ Q & \boldsymbol{0} \end{pmatrix} 2^{-sp}\right] = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}\,c \end{pmatrix} \exp\left[\begin{pmatrix} \boldsymbol{0} & P\,c \\ Q/c & \boldsymbol{0} \end{pmatrix} 2^{-sp}\right] \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}/c \end{pmatrix} \tag{B.7}$$

for any non-zero c. Using this form of the exponential, it suffices to choose sp so that

$$\max[\left\| P\,c/2^{sp} \right\|_\infty, \left\| (Q/c)/2^{sp} \right\|_\infty] < 1/2 \tag{B.8}$$

c can be chosen to minimize the left side of equation B.8. The minimization condition is

$$\left\| P\,c \right\|_\infty = \left\| Q/c \right\|_\infty \quad \to \quad c = \sqrt{\left\| Q \right\|_\infty / \left\| P \right\|_\infty} \tag{B.9}$$

which reduces equation B.8 to

$$\sqrt{\left\| P \right\|_\infty \left\| Q \right\|_\infty} / 2^{sp} < 1/2 \tag{B.10}$$

The scaling power sp defined by equation B.10 is typically significantly less than that defined by equation B.5.

After determining sp to satisfy equation B.10, the scaling factor is applied to A,

$$A \leftarrow A/2^{sp}, \quad P \leftarrow P/2^{sp}, \quad Q \leftarrow Q/2^{sp} \tag{B.11}$$

The exponential of A is then computed using a (6,6) Padé approximation,

$$\exp[A] \cong \left(\sum_{j=0}^{6} c_j \, (-A)^j \right)^{-1} \left(\sum_{j=0}^{6} c_j \, A^j \right) \tag{B.12}$$

wherein the c_j coefficients are

$$c_j = \frac{(12-j)!\,6!}{12!\,j!\,(6-j)!} \tag{B.13}$$

The coefficients' numerical values are

$$\{c_0, c_1, c_2, c_3, c_4, c_5, c_6\} = \{1, \frac{1}{2}, \frac{5}{44}, \frac{1}{66}, \frac{1}{792}, \frac{1}{15840}, \frac{1}{665280}\} \tag{B.14}$$

Substituting equation B.2, equation B.12 becomes

$$\exp\left[\begin{pmatrix} \boldsymbol{0} & P \\ Q & \boldsymbol{0} \end{pmatrix}\right] \cong \begin{pmatrix} B00 & -B01 \\ -B10 & B11 \end{pmatrix}^{-1} \begin{pmatrix} B00 & B01 \\ B10 & B11 \end{pmatrix} \tag{B.15}$$

wherein

$$B00 = c_0 \, \boldsymbol{I} + (c_2 \, \boldsymbol{I} + (c_4 \, \boldsymbol{I} + c_6 \, P Q) \, P Q) \, P Q \tag{B.16}$$

$$B11 = c_0 \, \boldsymbol{I} + Q \, (c_2 \, \boldsymbol{I} + (c_4 \, \boldsymbol{I} + c_6 \, P Q) \, P Q) \, P \tag{B.17}$$

$$B01 = (c_1 \, \boldsymbol{I} + (c_3 \, \boldsymbol{I} + c_5 \, P Q) \, P Q) \, P \tag{B.18}$$

$$B10 = Q \, (c_1 \, \boldsymbol{I} + (c_3 \, \boldsymbol{I} + c_5 \, P Q) \, P Q) \tag{B.19}$$

Note the common subexpressions in the above equations. As a consequence of this commonality, the following relationships hold,

$$B00 \; B01 = B01 \; B11 \tag{B.20}$$

$$B11 \; B10 = B10 \; B00 \tag{B.21}$$

from which it follows that equation B.15 can be equivalently stated,

$$\exp\left[\begin{pmatrix} \boldsymbol{0} & P \\ Q & \boldsymbol{0} \end{pmatrix}\right] \cong$$

$$\begin{pmatrix} B00 & B01 \\ B10 & B11 \end{pmatrix} \begin{pmatrix} (B00^2 - B01\,B10)^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & (B11^2 - B10\,B01)^{-1} \end{pmatrix} \begin{pmatrix} B00 & B01 \\ B10 & B11 \end{pmatrix} \tag{B.22}$$

Also, the special form of P in equations 13.112 and 14.22, as defined by equations 13.108 and 14.12, can be exploited to optimize the calculation of $B00$, $B01$, $B10$, $B11$.

Appendix C. Derivation of the symmetry relations 7.3

The stratum symmetry relations 7.3 apply generally when the S matrix characterizes a grating region that is invariant under an x_1 coordinate transformation that maps $x_1^{[0]}$ to $x_1^{[1]}$ and $x_1^{[1]}$ to $x_1^{[0]}$ (cf. equation 7.1). The symmetry can be demonstrated directly from the S matrix construction procedure outlined in section 13, as follows.

The exponential matrix in equation 13.112 satisfies the following condition,

$$\begin{pmatrix} \Phi EE & \Phi EH \\ \Phi HE & \Phi HH \end{pmatrix} \begin{pmatrix} \Phi EE & -\Phi EH \\ -\Phi HE & \Phi HH \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \tag{C.1}$$

This is a result of the zero blocks in the exponential argument,

$$\begin{aligned} \begin{pmatrix} \Phi EE & -\Phi EH \\ -\Phi HE & \Phi HH \end{pmatrix} &= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix} \begin{pmatrix} \Phi EE & \Phi EH \\ \Phi HE & \Phi HH \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix} \exp\left[\begin{pmatrix} \boldsymbol{0} & DEH \\ -DHE & \boldsymbol{0} \end{pmatrix} \Delta x_1\right] \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix} \\ &= \exp\left[\begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{0} & DEH \\ -DHE & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix} \Delta x_1\right] \\ &= \exp\left[\begin{pmatrix} \boldsymbol{0} & -DEH \\ DHE & \boldsymbol{0} \end{pmatrix} \Delta x_1\right] = \begin{pmatrix} \Phi EE & \Phi EH \\ \Phi HE & \Phi HH \end{pmatrix}^{-1} \end{aligned} \tag{C.2}$$

Equation 13.117 has the form

$$\Phi^{[\pm\pm']} = \tfrac{1}{2}\left(A^{-1}\left(\Phi EE\,A \pm' \Phi EH\,B\right) \pm B^{-1}\left(\Phi HE\,A \pm' \Phi HH\,B\right)\right) \tag{C.3}$$

wherein

$$A = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & (\lambda / \varepsilon^{[f]})\, df_1^{[+,j]} \end{pmatrix} \tag{C.4}$$

$$B = i \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\lambda\, df_1^{[+,j]} & \boldsymbol{0} \end{pmatrix} \tag{C.5}$$

The different sign combinations in equation C.3 yield the following expression,

$$\begin{pmatrix} \Phi^{[++]} & \Phi^{[+-]} \\ \Phi^{[-+]} & \Phi^{[--]} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} A^{-1} & B^{-1} \\ A^{-1} & -B^{-1} \end{pmatrix} \begin{pmatrix} \Phi EE & \Phi EH \\ \Phi HE & \Phi HH \end{pmatrix} \begin{pmatrix} A & A \\ B & -B \end{pmatrix} \tag{C.6}$$

This relationship can be inverted to obtain,

$$\begin{pmatrix} \Phi EE & \Phi EH \\ \Phi HE & \Phi HH \end{pmatrix} = \frac{1}{2} \begin{pmatrix} A & A \\ B & -B \end{pmatrix} \begin{pmatrix} \Phi^{[++]} & \Phi^{[+-]} \\ \Phi^{[-+]} & \Phi^{[--]} \end{pmatrix} \begin{pmatrix} A^{-1} & B^{-1} \\ A^{-1} & -B^{-1} \end{pmatrix} \tag{C.7}$$

Substitution of equation C.7 in C.1 yields

$$\begin{pmatrix} \Phi^{[++]} & \Phi^{[+-]} \\ \Phi^{[-+]} & \Phi^{[--]} \end{pmatrix} \begin{pmatrix} \Phi^{[--]} & \Phi^{[-+]} \\ \Phi^{[+-]} & \Phi^{[++]} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \tag{C.8}$$

The left two blocks in this expression are

$$\Phi^{[++]}\, \Phi^{[--]} + \Phi^{[+-]}\, \Phi^{[+-]} = I \tag{C.9}$$

$$\Phi^{[-+]}\, \Phi^{[--]} + \Phi^{[--]}\, \Phi^{[+-]} = 0 \tag{C.10}$$

These two identities imply the following,

$$-(\Phi^{[--]})^{-1}\, \Phi^{[-+]} = \Phi^{[+-]}\, (\Phi^{[--]})^{-1} \tag{C.11}$$

$$\Phi^{[++]} - \Phi^{[+-]}\, (\Phi^{[--]})^{-1}\, \Phi^{[-+]} = (\Phi^{[--]})^{-1} \tag{C.12}$$

Comparing equations C.11 and C.12 with equation 13.120, the symmetry relations 7.3 are obtained.

Appendix D. Computation of the field inside the grating

The electromagnetic field inside the grating can be determined by applying equations 8.1 and 8.2 wherein $x_1^{[0]}$ and $x_1^{[2]}$ are the x_1 coordinates at the bottom and top of the grating ($x_1^{[0]} = b_1[0]-0$, $x_1^{[2]} = b_1[L_1]+0$; cf. equations 3.1 and 3.2). $x_1^{[1]}$ is the x_1 level at which the field is to be determined, and Sa and Sb are the cumulative S matrices for the portions of the grating below and above $x_1^{[1]}$, respectively.

The incident field below the grating ($F^{[+]}[x_1^{[0]}]$) is zero, so equation 8.1 reduces to

$$\begin{pmatrix} F^{[-]}[x_1^{[0]}] \\ F^{[+]}[x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} Sa01 \\ Sa11 \end{pmatrix} F^{[-]}[x_1^{[1]}] \tag{D.1}$$

In principle, the transmitted field ($F^{[-]}[x_1^{[0]}]$) could be computed and used in equation D.1 to determine the field at $x_1^{[1]}$ ($F^{[-]}[x_1^{[1]}] = Sa01^{-1}\ F^{[-]}[x_1^{[0]}]$, $F^{[+]}[x_1^{[1]}] = Sa11\, F^{[-]}[x_1^{[1]}]$). However, this process could be numerically unstable if, for example, the lower portion of the grating is opaque so that $Sa01$ and $F^{[-]}[x_1^{[0]}]$ are both zero. This difficulty can be avoided by determining $F^{[-]}[x_1^{[1]}]$ from equation 8.2,

$$F^{[-]}[x_1^{[1]}] = Sb00\, F^{[+]}[x_1^{[1]}] + Sb01\, F^{[-]}[x_1^{[2]}] \tag{D.2}$$

The $F^{[-]}[x_1^{[2]}]$ term in equation D.2 represents the incident field, and the $F^{[+]}[x_1^{[1]}]$ term can be determined by substituting equation D.2 in D.1,

$$F^{[+]}[x_1^{[1]}] = (\boldsymbol{I} - Sa11\, Sb00)^{-1}\ Sa11\, Sb01\, F^{[-]}[x_1^{[2]}] \tag{D.3}$$

Thus, the internal field has the following form,

$$\begin{pmatrix} F^{[+]}[x_1^{[1]}] \\ F^{[-]}[x_1^{[1]}] \end{pmatrix} = \begin{pmatrix} S^{[+]} \\ S^{[-]} \end{pmatrix} F^{[-]}[x_1^{[2]}] \tag{D.4}$$

wherein

$$S^{[+]} = (\boldsymbol{I} - Sa11\, Sb00)^{-1}\ Sa11\, Sb01 \tag{D.5}$$

$$S^{[-]} = Sb00\ S^{[+]} + Sb01 \tag{D.6}$$

The computation procedure defined by equations D.5 and D.6 depends only on $Sa11$, $Sb00$, and $Sb01$. $Sa11$ can be computed efficiently using bottom-up stacking (because in the stacking equation 8.8, $S11$ does not depend on $Sa00$, $Sa01$ or $Sa10$);

and $Sb00$ and $Sb01$ can be computed efficiently by using top-down stacking (because in equation 8.8 $S00$ and $S01$ do not depend on $Sb10$ or $Sb11$).

After determining $F^{[+]}[x_1^{[1]}]$ and $F^{[-]}[x_1^{[1]}]$, the grating-tangential electromagnetic field projections can be determined from equations 6.50-53. If the grating is homogeneous at $x_1 = x_1^{[1]}$, the total electromagnetic field can then be determined by applying equations 6.48 and 6.49. (These equations determine the diffracted orders' $\hat{p}$ projections.) If the grating is inhomogeneous at $x_1 = x_1^{[1]}$, then the derivations in section 13 apply. Equations 13.94-97, which are equivalent to equations 6.50-53, determine the tangential field projections; and the grating-normal projections ($\mathit{ffH}_1^{[j]}$ and $\mathit{ffE}_1^{[j]}$) are determined by equations 13.74 and 13.80.

References

1. Michel Nevière, Evgeny Popov, *Light Propagation in Periodic Media*, (MarcelDekker, Inc. New York, 2003).

2. L. Li, "Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings," J. Opt. Soc. Am. A 13, 1024-1035, 1996.

3. L. Li, "Use of Fourier series in the analysis of discontinuous periodic structures," J. Opt. Soc. Am. A 13, 1870-1876, 1996.

4. L. Li, "New formulation of the Fourier modal method for crossed surface-relief gratings," J. Opt. Soc. Am. A 14, 2758-2769, 1997.

5. L. Li, "Multilayer modal method for diffraction gratings of arbitrary profile, depth, and permittivity," J. Opt. Soc. Am. A 10, 2581-2591, 1993.

6. C. Moler and C. Van Loan, "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later," SIAM Review, Vol. 45, No. 1 (2003), pp. 3-49.