Grating Diffraction Calculator (GD-Calc®) –

Demo and Tutorial Guide

Overview

GD-Calc is implemented in MATLAB® [1] (Version R2022a or later) and comprises the following files:

Documentation:
- GD-Calc_Intro.pdf: Tutorial introduction.
- GD-Calc_Demo.pdf: Description of the demo scripts.
- GD-Calc.pdf: Part 1 of GD-Calc.pdf explains the technical concepts, definitions and conventions upon which the software user interface is based. Part 2 details the theory and methods underlying the computation algorithms.

Core functionality:
- gdc.m: This is the entry point to the diffraction calculation engine. gdc.m performs data validation and calls gdc_engine to run the calculations. gdc.m can be called with no output arguments to just check data validity, or with no inputs or outputs to just display version and copyright information.
- gdc_engine.p: The calculation engine. gdc_engine is only intended to be called by gdc.

Basic utilities:
- gdc_plot.m: for plotting grating structures. (gdc_plot.m calls gdc.m to check data validity.)
- gdc_eff.m: This function converts the output of gdc.m into diffraction efficiency data.

Demo scripts:
- gdc_demo*.m: These demo scripts cover a variety of tutorial topics and demonstrate GD-Calc's computational capabilities. The code patterns in the demo scripts can adapted the for users' applications.

Demo utilities:
- circle_partition.m: This is a useful utility for defining circular grating structures such as posts and holes. circle_partition.m is used in gdc_demo5.m.
- Kogelnik.m: This function uses a simplified two-wave coupled wave theory to calculate a volume holographic grating's diffraction efficiency. Kogelnik.m is used by gdc_demo10.m.
- read_nk.m: This function reads refractive index files of the type illustrated by the *.nk files. (See the comment header in read_nk.m for information on where to get nk files for common materials.) read_nk.m is used in gdc_demo11.m.
- EH_map.m: This function converts the gdc output data into a spatial map of the internal electromagnetic field within a grating structure. EH_map.m is used by gdc_demo13.m, gdc_demo14.m, and gdc_demo15.m.

---

[1] MATLAB is a registered trademarks of The MathWorks, Inc. (www.mathworks.com).

The above files can be downloaded from CodeOcean (https://codeocean.com/explore – search for "GD-Calc" to find the latest version). With the exception of the *.pdf, *.p and gdc.m files, all are public domain.

Each demo script contains a comment header describing briefly the type of grating structure modeled and what tutorial topics are covered by the example. The demo scripts were tested on a 6-core, 3.6 GHz Intel Xeon system with 128 GB RAM. The runtime for most demos is at most a few seconds (except gdc_demo14.m, which takes 30 seconds). The runtime and memory requirements are primarily determined by the **m_max** parameter (diffraction order truncation limit). For biperiodic gratings, memory requirements typically scale in approximate proportion to $\texttt{m\_max}^4$, and runtime scales in approximate proportion to $\texttt{m\_max}^6$. In practice, some experimentation might be needed to determine the order truncation limit that best balances the tradeoff between convergence (with respect to **m_max**) and computational resource requirements (time and memory).

Table 1 briefly outlines the topics covered by the demo scripts. The first few demo scripts (1a, 1b, 1c, and 2-8) represent test cases from the published literature on grating diffraction theory. The demo scripts 9-12 are representative of interesting grating designs that have practical engineering applications (an alignment sensor, a volume Bragg-diffraction grating, a photonic crystal, and an achromatizing, blazed transmission grating). Demo scripts 13-15 illustrate the calculation and display of internal electromagnetic fields. Demo 16 models a coherent beam combiner with multiple, plane-wave incident beams. Demo 17 models diffraction of a Gaussian incident beam represented as a Fourier series of incident plane waves.

Table 1: GD-Calc demo scripts

demo:	1a	1b	1c	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
compare to published data	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓									
uniperiodic grating	✓	✓	✓								✓	✓		✓	✓		✓	✓	✓
biperiodic grating				✓	✓	✓	✓	✓	✓	✓			✓			✓			
circular structure							✓												
multilayer grating											✓				✓				
polarization geometry	✓				✓														
choice of unit cell					✓	✓		✓	✓										
non-orthogonal unit cell							✓	✓	✓										
diffraction order selection					✓	✓	✓		✓										
accuracy limitations			✓		✓	✓	✓	✓	✓	✓						✓			
harmonic indices							✓	✓	✓		✓		✓		✓				
parameterization											✓	✓			✓				
coordinate break											✓	✓	✓		✓				
replication module											✓	✓							
refractive index tables														✓					
internal fields															✓	✓	✓		
animation											✓				✓	✓			
multiple incident orders																		✓	✓
Gaussian incident beam																			✓

Demo 1a: Uniperiodic, sinusoidal grating

This demo computes multi-order diffraction efficiencies of the sinusoidal grating structure illustrated in Figure 1. For the purpose of illustration, Figure 1 shows the structure as being approximated by 10 stacked, lamellar grating strata, although demo 1a actually uses 50 strata. For consistency with the literature, each stratum's wall positions are defined so that the sinusoidal profile bisects each wall at its vertical midpoint. (A slightly more accurate approach would be to set each wall position according the sinusoid's average horizontal intercept within the stratum. To use this option set the **ctr_sect** toggle to **false**.)
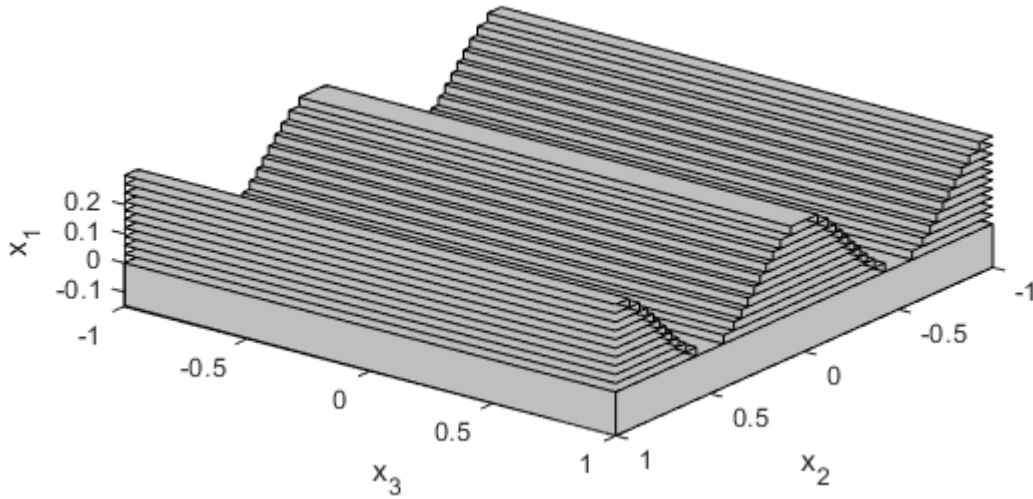

Figure 1. Demo 1a grating structure (as depicted by gdc_plot.m).

The grating geometry is fairly simple, but the incident beam's geometry and polarization state are somewhat complex in this example. The beam direction is defined in terms of polar coordinates with the polar axis parallel to the grating lines, as illustrated in Figure 2. $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ are unit basis vectors corresponding to the x_1, x_2, and x_3 coordinate directions in Figure 1, and $\vec{f}$ is the incident field's wave vector (in spatial frequency units). The wave vector's direction is defined in terms of polar angle θ and azimuthal angle φ (respectively designated as **theta3** and **phi3** in the .m file) as illustrated in the figure,

$$\vec{f} = \tfrac{1}{\lambda}(-\sin[\theta]\cos[\varphi]\,\hat{e}_1 + \sin[\theta]\sin[\varphi]\,\hat{e}_2 + \cos[\theta]\,\hat{e}_3) \tag{1}$$

where λ is the wavelength. (The notational conventions used in Eq. 1 and throughout this document are outlined in GD-Calc.pdf, Section 2.)
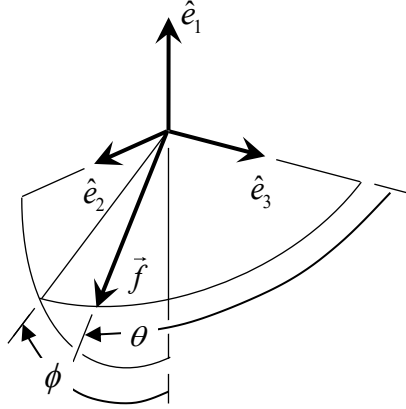
Figure 2. Incident beam direction for demo1a.

The incident electric field amplitude $\vec{E}$ is specified in relation to two polarization basis vectors: a unit vector $\hat{s}$ that is orthogonal to $\hat{e}_1$ and $\vec{f}$ (i.e., orthogonal to the incidence plane), and a unit vector $\hat{p}$ that is orthogonal to $\hat{s}$ and $\vec{f}$ (i.e., parallel to the incidence plane; cf. GD-Calc.pdf, Eq's. 4.17-20). The three vectors $\hat{s}$, $\hat{p}$, and $\lambda \vec{f}$ form an orthonormal basis set with

$$\hat{p} = \lambda \vec{f} \times \hat{s} \tag{2}$$

Denoting the incident $\vec{E}$ field's $\hat{s}$ and $\hat{p}$ projections as A and B, the field has the form

$$\vec{E} = A\hat{s} + B\hat{p} \tag{3}$$

The incident $\vec{H}$ field is

$$\vec{H} = \lambda \vec{f} \times \vec{E} = A\hat{p} - B\hat{s} \tag{4}$$

(cf. GD-Calc.pdf, Eq. 6.11).

Two polarization states are considered in demo 1a: one with the incident $\vec{H}$ field linearly polarized orthogonal to $\hat{e}_3$ (the grating line direction), and one with the incident $\vec{E}$ field orthogonal to $\hat{e}_3$. For the second case ($E_3 = 0$), the A and B amplitudes are defined by the following condition in which C is an arbitrary constant,

$$A = -(\hat{e}_3 \bullet \hat{p})\, C, \quad B = (\hat{e}_3 \bullet \hat{s})\, C \quad \rightarrow \quad \hat{e}_3 \bullet \vec{E} = 0 \tag{5}$$

Similarly for the first case ($H_3 = 0$), the following condition applies,

$$A = (\hat{e}_3 \bullet \hat{s})\, C, \quad B = (\hat{e}_3 \bullet \hat{p})\, C \quad \rightarrow \quad \hat{e}_3 \bullet \vec{H} = 0 \tag{6}$$

In gdc_demo1a.m the constant C is defined so that $|A|^2 + |B|^2 = 1$.

gdc.m does not return diffraction efficiencies directly, but it returns amplitude transmission and reflectance matrices from which the efficiencies can be calculated. The utility function gdc_eff.m takes the output from gdc.m to perform this calculation. The efficiencies are calculated for four particular incident polarizations: $\vec{E} = \hat{s}$, $\hat{p}$, $(\hat{s} + \hat{p})/\sqrt{2}$, and $(\hat{s} - i\,\hat{p})/\sqrt{2}$, and the corresponding efficiencies in the gdc_eff.m output are denoted as **eff1**, **eff2**, **eff3** and **eff4**. The function output includes these efficiency values for each transmitted and reflected diffraction order. Table 2 lists the four polarization states and corresponding efficiencies. The four efficiency values can also be combined to calculate the diffraction efficiency for an arbitrary incident polarization state, as indicated in the last row of Table 2. The function **eff**$[A, B]$ has the form,

$$\textbf{eff}[A,B] = \frac{\left(\begin{array}{l} \textbf{eff1}\,|A|^2 + \textbf{eff2}\,|B|^2 \\ \quad + (2\,\textbf{eff3} - \textbf{eff1} - \textbf{eff2})\,\mathrm{Re}[A\,B^*] + (2\,\textbf{eff4} - \textbf{eff1} - \textbf{eff2})\,\mathrm{Im}[A\,B^*] \end{array}\right)}{|A|^2 + |B|^2} \tag{7}$$

(GD-Calc.pdf, Eq's. A.28 and A.30). This calculation is performed in gdc_demo1a.m for each of the two polarizations defined by Eq's. 5 and 6 above.

Table 2. Dependence of diffraction efficiency on incident polarization state.

Incident polarization	Diffraction efficiency
$\vec{E} = \hat{s}$	**eff1**
$\vec{E} = \hat{p}$	**eff2**
$\vec{E} = (\hat{s} + \hat{p})/\sqrt{2}$	**eff3**
$\vec{E} = (\hat{s} - i\,\hat{p})/\sqrt{2}$	**eff4**
$\vec{E} = A\hat{s} + B\hat{p}$	**eff**$[A,B]$

Demo 1b, 1c: Uniperiodic, sinusoidal grating (deep profile, metallic)

Demo 1b and 1c are similar to demo 1a, but demo 1b models a deep dielectric grating, and demo 1c models a deep metallic grating. The incident field specifications in demo 1b and 1c are modified to model planar (non-conical) diffraction, i.e., the incident wave vector is orthogonal to the grating lines. Polar angle coordinates are defined with the polar axis normal to the grating substrate, not parallel to the grating lines as in demo 1a.

Demo 1b and 1c replicate published test cases showing the computation algorithm's good numerical stability on extremely deep grating structures; however, stability does not imply accuracy. Good accuracy can be easily achieved for TE polarization, but for deep structures with non-lamellar grating profiles – especially non-lamellar metallic structures – good TM accuracy is not as easily achieved. This is because the electric field exhibits large spikes in the corner regions of staircase-type profiles, and as the number of strata is increased, the spikes become narrower and the number of calculated diffraction orders must be increased in proportion to the stratum number to achieve adequate spatial sampling density across the spikes[2].

The difference between TE and TM accuracy performance is illustrated in Figures 3 and 4 for the metallic sinusoidal grating of demo 1c, with the grating height equal to the period. These figures show the computed zero-order reflectance (TE in Figure 3 and TM in Figure 4), plotted against the maximum diffraction order index (`m_max`) for each of several stratification numbers (`L1`). For TE polarization, convergence is quickly obtained with a small number of strata and orders. By contrast, TM polarization clearly requires a large number of strata, and a correspondingly large number of diffraction orders, to approach convergence. For challenging problems of this type, the computation results should not be accepted without first performing a two-parameter convergence test, varying both the stratification number and the order truncation limit.

---

[2] This phenomenon is discussed in section VI.5 of *Light Propagation in Periodic Media*, by Michel Nevière, Evgeny Popov (MarcelDekker, Inc. New York, 2003). Also see demo 15.
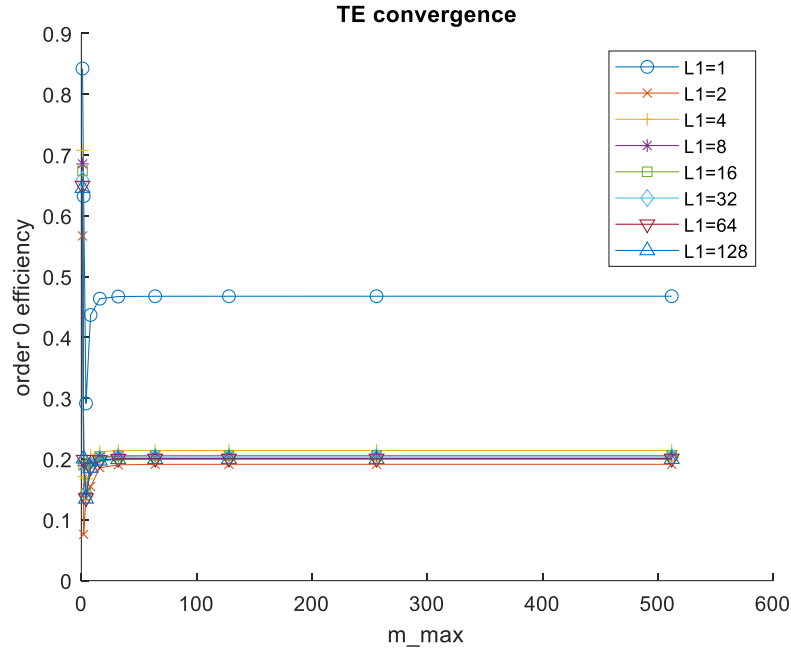
Figure 3. Zero-order TE convergence test for a sinusoidal, metallic grating with permittivity $(0.3 + 7.0i)^2$, period $d = 1$, wavelength $\lambda = d/1.7$, height $h = d$, and incidence angle $\theta = 30°$. The diffraction order truncation limit is **m_max**, and the number of strata is **L1**.
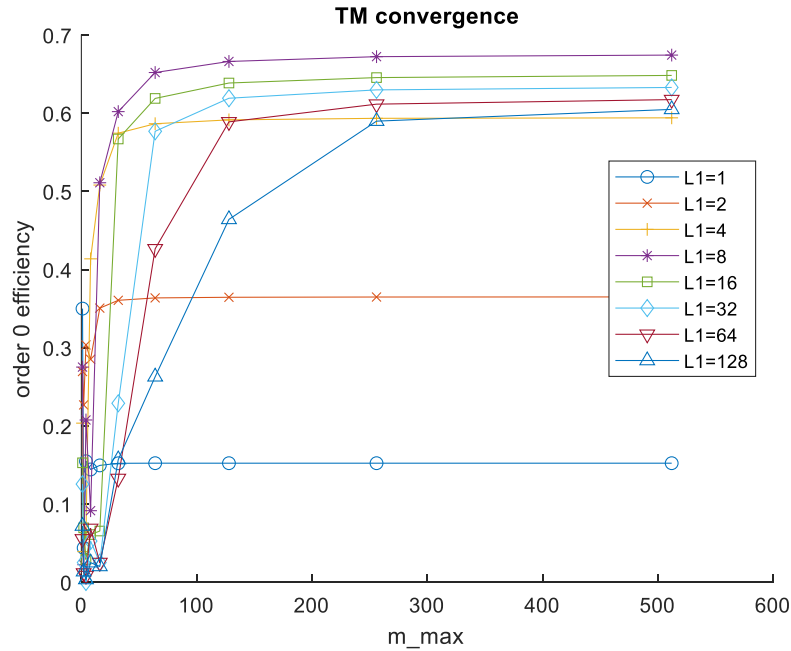

Figure 4. Same as Figure 3, but for TM polarization.

Demo 2: Biperiodic grating comprising rectangular pyramids

This demo illustrates the basic functionality of GD-Calc for modeling biperiodic gratings. The grating structure comprises rectangular pyramids, as illustrated in Figure 5.
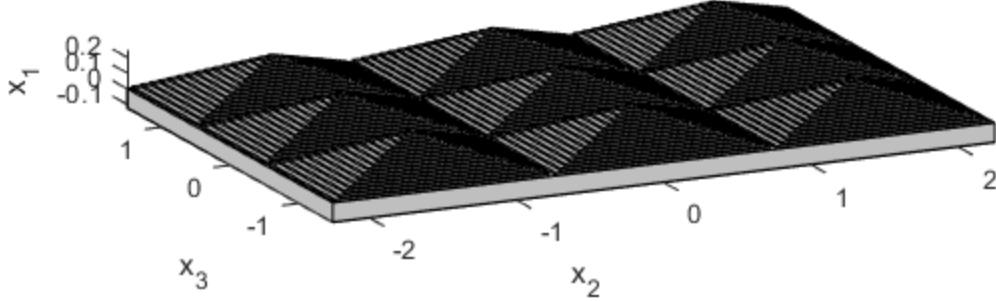

Figure 5: Biperiodic grating of demo 2.

Demo 3 and demo 4: Biperiodic checkerboard grating

Demo 3 and demo 4 model a checkerboard grating comprising square wells recessed in a dielectric substrate, as illustrated in Figure 6. Demo 3 uses unit cell B to define the grating periods, whereas demo 4 uses unit cell A. The grating is characterized by two period vectors, which are defined by the unit cell edges, and which are denoted as $\vec{d}_1^{[g,A]}$ and $\vec{d}_2^{[g,A]}$ for unit cell A and as $\vec{d}_1^{[g,B]}$ and $\vec{d}_2^{[g,B]}$ for unit cell B. The periods have the following coordinate representations,

$$\vec{d}_1^{[g,A]} = 2w\,\hat{e}_2 \tag{8}$$

$$\vec{d}_2^{[g,A]} = 2w\,\hat{e}_3 \tag{9}$$

$$\vec{d}_1^{[g,B]} = w(\hat{e}_2 + \hat{e}_3) \tag{10}$$

$$\vec{d}_2^{[g,B]} = w(\hat{e}_3 - \hat{e}_2) \tag{11}$$

where w is the checkerboard square width.

Each pair of grating periods defines an associated pair of fundamental grating frequencies. For example, the periods $\vec{d}_1^{[g,A]}$ and $\vec{d}_2^{[g,A]}$ define the associated frequencies $\vec{f}_1^{[g,A]}$ and $\vec{f}_2^{[g,A]}$ according to the reciprocal relationships,

$$\left.\begin{array}{ll} \vec{f}_1^{[g,A]} \bullet \vec{d}_1^{[g,A]} = 1, & \vec{f}_1^{[g,A]} \bullet \vec{d}_2^{[g,A]} = 0 \\ \vec{f}_2^{[g,A]} \bullet \vec{d}_1^{[g,A]} = 0, & \vec{f}_2^{[g,A]} \bullet \vec{d}_2^{[g,A]} = 1 \end{array}\right\} \tag{12}$$

(cf. GD-Calc.pdf, Eq. 3.29). The grating frequencies $\vec{f}_1^{[g,A]}$ and $\vec{f}_2^{[g,A]}$ defined by unit cell A, and the frequencies $\vec{f}_1^{[g,B]}$ and $\vec{f}_2^{[g,B]}$ defined by cell B, have the coordinate representations,

$$\vec{f}_1^{[g,A]} = \frac{1}{2w}\,\hat{e}_2 \tag{13}$$

$$\vec{f}_2^{[g,A]} = \frac{1}{2w}\,\hat{e}_3 \tag{14}$$

$$\vec{f}_1^{[g,B]} = \frac{1}{2w}(\hat{e}_2 + \hat{e}_3) = \vec{f}_1^{[g,A]} + \vec{f}_2^{[g,A]} \tag{15}$$

$$\vec{f}_2^{[g,B]} = \frac{1}{2w}(\hat{e}_3 - \hat{e}_2) = \vec{f}_2^{[g,A]} - \vec{f}_1^{[g,A]} \tag{16}$$
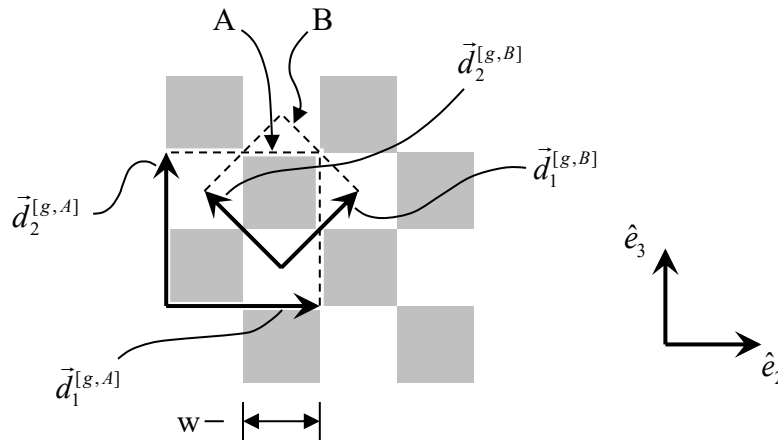

Figure 6. Checkerboard grating of demo 3 and demo 4.

The frequency relationships in Eq's. 15 and 16 can be used to determine the relationship between diffraction order indices with the two alternative unit cells. Each diffraction order has associated order indices $m_1^{[A]}$ and $m_2^{[A]}$ relative to unit cell A, such that the order's grating-tangential frequency differs from that of the incident wave by the increment $m_1^{[A]}\,\vec{f}_1^{[g,A]} + m_2^{[A]}\,\vec{f}_2^{[g,A]}$ (GD-Calc.pdf, Eq. 4.7). Similarly, the same order has associated order indices $m_1^{[B]}$ and $m_2^{[B]}$ relative to unit cell B such that the frequency difference is $m_1^{[B]}\,\vec{f}_1^{[g,B]} + m_2^{[B]}\,\vec{f}_2^{[g,B]}$. The relationship between the cell-A and cell-B indices is determined by the equivalence,

$$m_1^{[A]}\,\vec{f}_1^{[g,A]} + m_2^{[A]}\,\vec{f}_2^{[g,A]} = m_1^{[B]}\,\vec{f}_1^{[g,B]} + m_2^{[B]}\,\vec{f}_2^{[g,B]} \tag{17}$$

Substituting from the right sides of Eq's. 15 and 16 in Eq. 17 and matching the coefficients of $\vec{f}_1^{[g,A]}$ and $\vec{f}_2^{[g,A]}$ separately, the following relationships defining the cell-A order indices from the cell-B indices are obtained,

$$m_1^{[A]} = m_1^{[B]} - m_2^{[B]} \tag{18}$$

$$m_2^{[A]} = m_1^{[B]} + m_2^{[B]} \tag{19}$$

Note that the above equations relating the order indices have the same form as the relationships between the grating periods,

$$\vec{d}_1^{[g,A]} = \vec{d}_2^{[g,B]} - \vec{d}_2^{[g,B]} \tag{20}$$

$$\vec{d}_2^{[g,A]} = \vec{d}_2^{[g,B]} + \vec{d}_2^{[g,B]} \tag{21}$$

This is true in general: Once the period relationships have been defined, the order index relationships are obtained by simply substituting m's for d's.

The cell-B order indices can similarly be obtained from the cell-A indices,

$$m_1^{[B]} = (m_1^{[A]} + m_2^{[A]})/2 \tag{22}$$

$$m_2^{[B]} = (m_2^{[A]} - m_1^{[A]})/2 \tag{23}$$

Some cell-A order indices $m_1^{[A]}$ and $m_2^{[A]}$ yield non-integer values of $m_1^{[B]}$ or $m_2^{[B]}$ in Eq's. 22 and 23. Based on the grating's periodic symmetry, these orders have exactly zero amplitude, so there is no need to include them in the diffraction calculation. Furthermore, for consistency the diffraction calculation with unit cell A or B should be done with the same set of diffraction orders, using Eq's. 18 and 19 to select the orders that are used in the cell-A calculation. This order selection process is illustrated by demo 3 and demo 4: In demo 3, the order indices $m_1^{[B]}$ and $m_2^{[B]}$ are simply defined to be all integers not exceeding **m_max** in magnitude. In demo 4, this same set of cell-B indices is used, and Eq's. 18 and 19 are used to define the cell-A indices. After the diffraction efficiency computation has been done, Eq's. 22 and 23 are used to convert the order indices back to their cell-B-relative values, and the diffraction orders are sorted to match their sequence in the demo 3 output.

The results of demo 3 and demo 4 also must be adjusted to ensure that they represent the same polarization state. As noted in the preceding demo 1a notes, the diffraction efficiencies computed by **gdc_eff** are defined relative to $\hat{s}$ and $\hat{p}$ polarization states, where the incident wave's $\hat{s}$ vector is orthogonal to the plane of incidence. But demo 3 and demo 4 assume normal incidence, in which case there is no well-defined plane of incidence. In this case, the $\hat{s}$ direction is defined to be orthogonal

to the grating's first frequency vector (GD-Calc.pdf, Eq. 4.20, "if $\ldots = 0$" branch). The first grating frequency for unit cell A ($\vec{f}_1^{[g,A]}$) differs from that of unit cell B ($\vec{f}_1^{[g,B]}$), so the gdc_eff output data for demo 3 and demo 4 correspond to different polarization states.

The output of both demo scripts is standardized to common polarization states by computing the diffraction efficiencies with the incident wave polarized parallel to $\hat{e}_2$ or $\hat{e}_3$. For normal incidence $\hat{s}$ and $\hat{p}$ are both parallel to the grating substrate, so $\hat{e}_2$ and $\hat{e}_3$ can be projected onto these vectors as follows,

$$\hat{e}_2 = (\hat{e}_2 \bullet \hat{s})\hat{s} + (\hat{e}_2 \bullet \hat{p})\hat{p} \tag{24}$$

$$\hat{e}_3 = (\hat{e}_3 \bullet \hat{s})\hat{s} + (\hat{e}_3 \bullet \hat{p})\hat{p} \tag{25}$$

Assuming an incident $\vec{E}$ field that is a scalar multiple of either of the polarization vectors represented by Eq's. 24 or 25, the corresponding diffraction efficiencies can be determined using Eq. 7. For unit cell A (demo 4) the $\hat{s}$ and $\hat{p}$ vectors are equal to $\hat{e}_3$ and $\hat{e}_2$, respectively (based on Eq. 13 above and GD-Calc.pdf, Eq's. 4.19 and 4.20), so the efficiencies for $\hat{e}_2$ and $\hat{e}_3$ polarization simply correspond respectively to `eff2` and `eff1`.

In setting up a biperiodic grating geometry, each biperiodic stratum is defined by partitioning the stratum into stripes parallel to the stratum's second period vector ($\vec{d}_2^{[g,B]}$ in demo 3, and $\vec{d}_2^{[g,A]}$ in demo 4), and then partitioning each stripe into homogeneous, rectangular structural blocks. This type of block-structured geometry naturally fits the grating configuration with unit cell A, as illustrated in Figure 7; but the diagonal orientation of unit cell B requires that the grating be partitioned into a large number of stripes that approximate the grating walls by sawtooth profiles, as illustrated in Figure 8. (For the purpose of illustration 10 stripes per period are shown in Figure 8, although demo3 actually uses 26 and the JOSA reference cited in demo 3 uses 1000. The light-shaded blocks in Figure 8 straddle the grating walls, and the permittivity within each of these blocks is defined to be the geometric mean of the grating's actual permittivity within the block.)
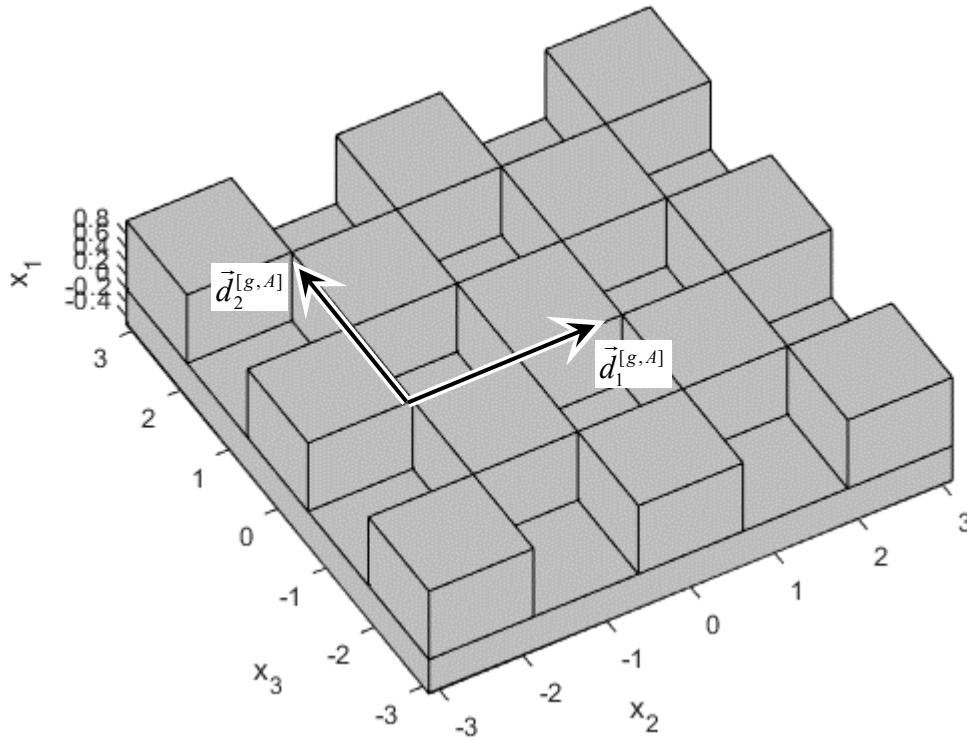
Figure 7. Checkerboard grating model with unit cell A (demo 4).
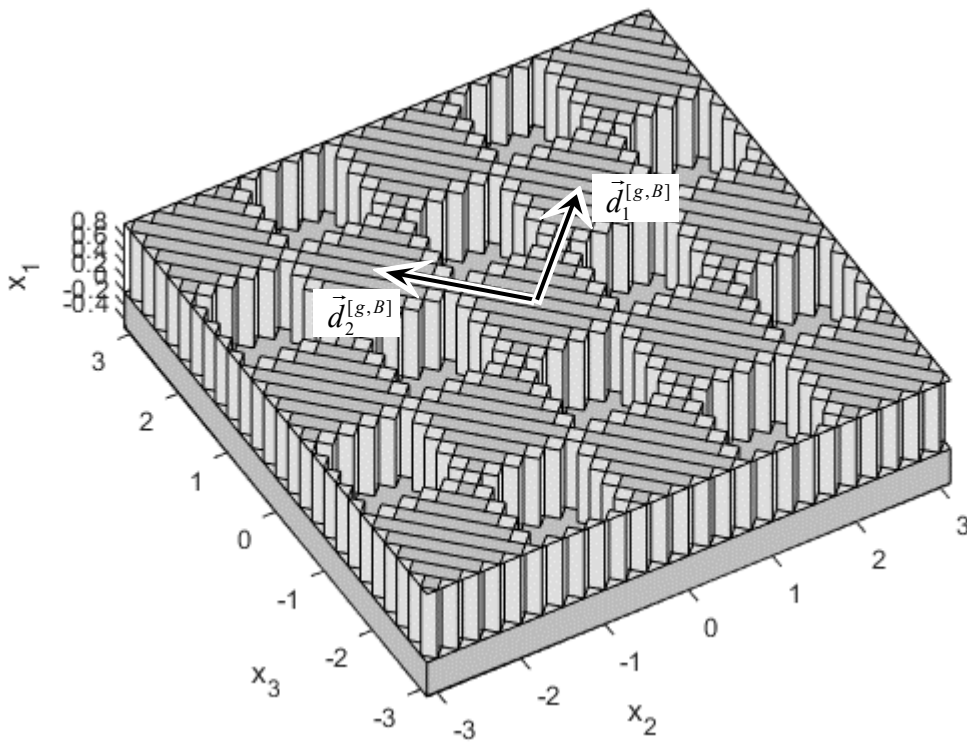

Figure 8. Checkerboard grating model with unit cell B (demo 3).

The stripe partitioning of grating strata introduces an artificial asymmetry into the mathematical model of the grating, which does not exist in the physical structure; and this asymmetry is reflected in the calculated diffraction efficiencies. For example, one would expect that under normal incidence the order-(1, 0) efficiency for either $\hat{e}_2$ or $\hat{e}_3$ polarization (using cell-B order indexing, ($m_1^{[B]}, m_2^{[B]}$)) would match the order-(0, 1) efficiency for the same polarization, which it does with unit cell A but not with cell B. On the other hand, one would also expect that the order-(1, 0) efficiency would be the same for both polarizations, which it is with cell B but not with cell A. However, the cell-A symmetry error is an order of magnitude smaller than that of cell B because the grating geometry naturally conforms to the block-structured model description with cell A. Thus, while it is generally desirable to use a unit cell of minimum area (e.g., cell B) to define the diffraction order selection, it may be advantageous to define the grating structure using a different unit cell (e.g. cell A) that requires very few structural blocks for the geometry specification and does not rely on the sawtooth wall approximation.

Figure 9 shows a graph of the order-(1, 0) and (0, 1) efficiencies, computed for a range order truncation limits **m_max**. (The total number of diffraction orders is $(2\,\mathbf{m\_max}+1)^2$.) Convergence lines are shown for several values of **L1**, the number of cell-B stripes per period. With $\mathbf{m\_max}=12$ the cell-A symmetry error is $5.5\cdot10^{-5}$. The cell-B data exhibits no significant symmetry error if only 10 stripes are used, but the convergent result is clearly in error (by $8.5\cdot10^{-4}$) relative to the cell-A efficiency value. Using 100 stripes, the cell-B result is much more consistent with cell A, but there is a symmetry error of $3.7\cdot10^{-4}$. Increasing the number of stripes to 1000 without also increasing **m_max** does not provide any benefit – it only increases the symmetry error (to $9.9\cdot10^{-4}$). This is clearly a manifestation of the same kind of problem that occurs when non-lamellar profiles are modeled with high stratification numbers: The electromagnetic field exhibits large spikes inside the "teeth" of the cell-B grating walls, and the field's Fourier expansion does not have sufficient spatial sampling density to resolve these spikes unless **m_max** is increased in proportion to **L2**.

The symmetry errors can be significantly affected by the diffraction order truncation method. The above example uses simple rectangular truncation, in which $m_1^{[B]}$ and $m_2^{[B]}$ both range from **−m_max** to **+m_max**. Figure 10 illustrates the symmetry error using an alternative "diagonal truncation" method in which the sum of $|m_1^{[B]}|$ and $|m_2^{[B]}|$ is limited by **m_max**. With this method, the convergence is less monotonic but the cell-B symmetry error with **L2** = 1000 is greatly reduced (and is virtually eliminated when **m_max** is odd). Also, for a specified **m_max** limit the total number of diffraction orders is reduced by almost a factor of 2 by using diagonal truncation, cutting computation time by about a factor of 8. (The "**alt_order**" flags in gdc_demo3.m and gdc_demo4.m can be toggled to use this alternate order truncation.)
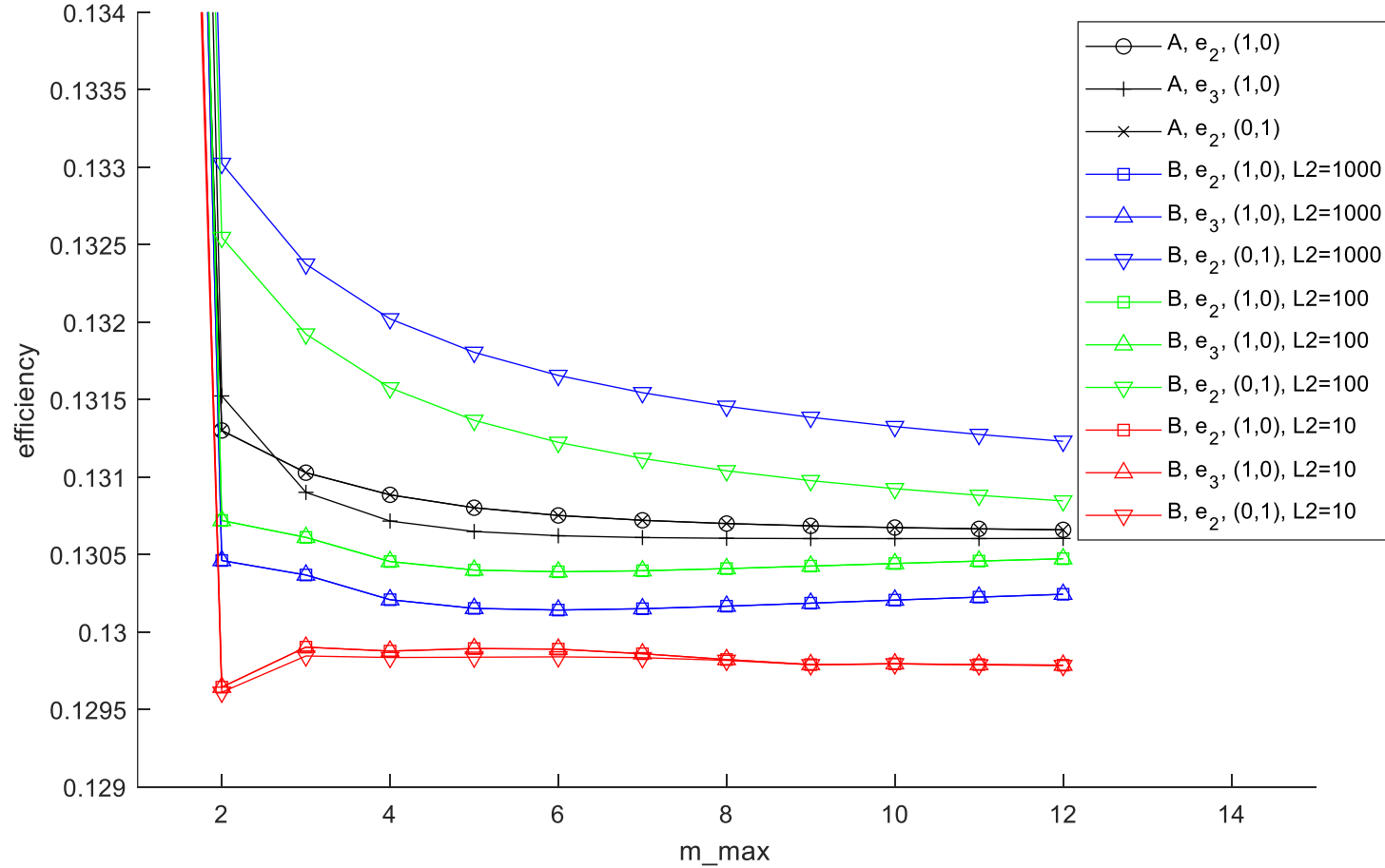
Figure 9. Convergence and stripe-induced symmetry error for checkerboard grating. The legend indicates the unit cell (A or B), the incident polarization direction ($\hat{e}_2$ or $\hat{e}_3$), the diffraction order ($m_1^{[B]}, m_2^{[B]}$) relative to cell B, and the number of cell-B stripes per period (**L2**). Rectangular order truncation is used ($|m_1^{[B]}| \leq$ **m_max**, $|m_2^{[B]}| \leq$ **m_max**).
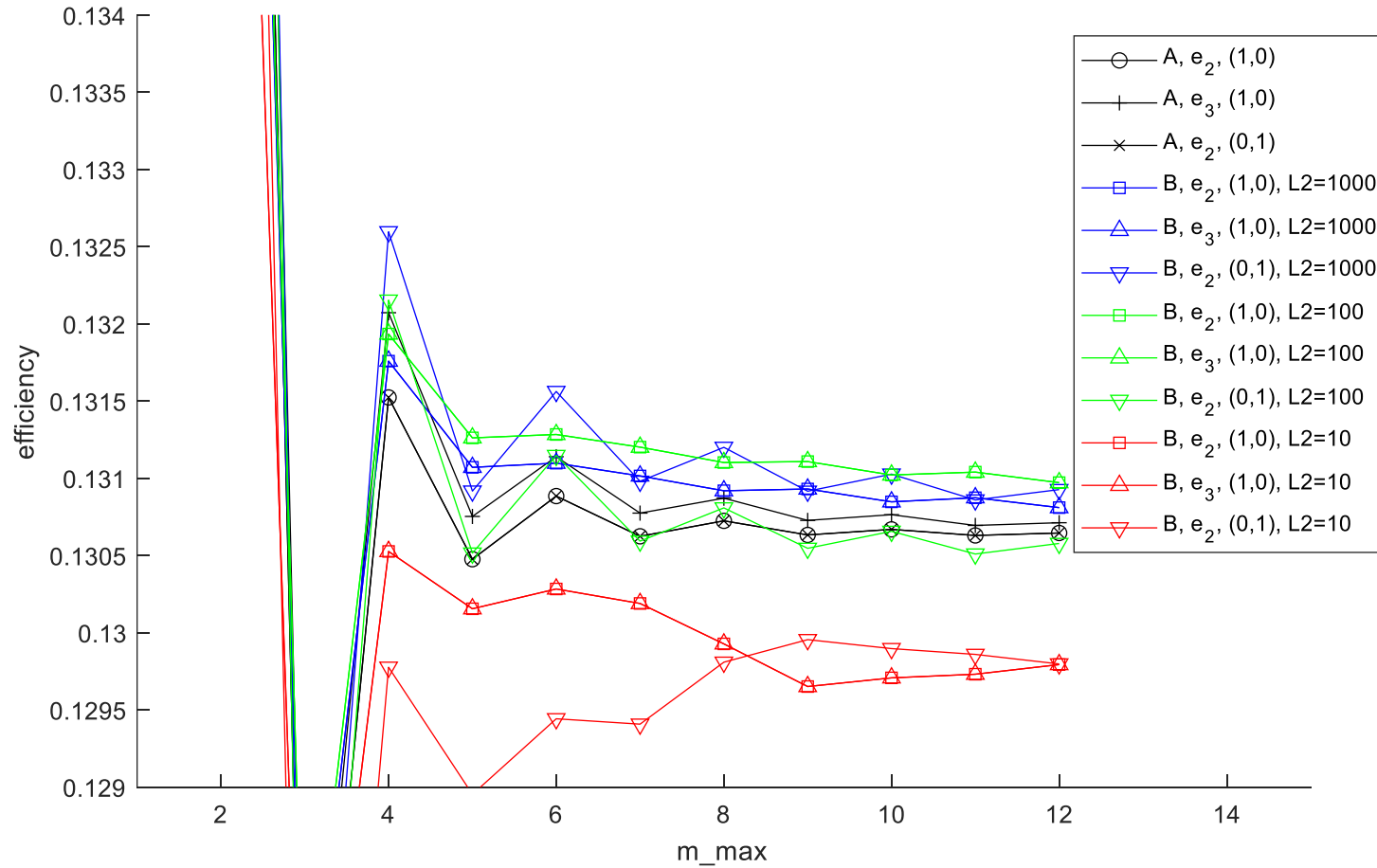
Figure 10. Same as Figure 9, but using diagonal order truncation ($|m_1^{[B]}| + |m_2^{[B]}| \leq$ **m_max**).

Demo 5: Biperiodic grating comprising circular pillars

Demo 5 models a grating structure comprising circular pillars centered at the points of an equilateral triangle array as illustrated in Figure 11. For the purpose of illustration Figure 11 shows nine stripes per pillar, although demo 5 actually uses 49 and the cited JOSA reference uses 1999. (The block partitioning of the circular structures is done with the aid of the utility function circle_partition.m, which is documented in Appendix A; see GD-Calc.pdf, Figure 4.) The grating's period vectors are designated as $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$ in Figure 11. The plane of incidence is parallel to the vector $\vec{d}_1^{[g]} + \vec{d}_2^{[g]}$, and the incident beam is linearly polarized in the plane of incidence, so the diffraction geometry has reflective symmetry across the incidence plane.

The diffraction orders are selected using a hexagonal order truncation scheme of the type illustrated in GD-Calc.pdf, Figure 5 and Eq's. 4.14 and 4.15. If simple rectangular truncation were to be used the selected orders would depend on the choice of unit cell. For example, with the unit cell defined by the basis vectors $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$ the index orders (m_1, m_2) would be defined by the truncation conditions $|m_1| \leq$ **m_max**, $|m_2| \leq$ **m_max**. Using the cell defined by the basis vectors $\vec{d}_1^{[g]}$ and $\vec{d}_1^{[g]} - \vec{d}_2^{[g]}$ the truncation conditions would be $|m_1| \leq$ **m_max**, $|m_1 - m_2| \leq$ **m_max**; and with basis vectors $\vec{d}_2^{[g]}$ and $\vec{d}_1^{[g]} - \vec{d}_2^{[g]}$ they would be $|m_2| \leq$ **m_max**, $|m_1 - m_2| \leq$ **m_max**. Demo 5 uses the intersection of these three selection sets to limit (m_1, m_2), i.e., the truncation conditions are $|m_1| \leq$ **m_max**, $|m_2| \leq$ **m_max**, $|m_1 - m_2| \leq$ **m_max**.

As with demo 3 and demo 4 the stripe partitioning of the pillars creates small numerical inaccuracies, as evidenced by asymmetries in the calculated diffraction efficiencies. For example, the calculated efficiency for the (-1, -2) transmitted order differs between the three alternative stripe orientations illustrated in Figures 11, 12, and 13. (In Figure 11 the stripes are parallel to $\vec{d}_2^{[g]}$, and in Figure 12 they are parallel to $\vec{d}_1^{[g]}$, and in Figure 13 they are parallel to $\vec{d}_1^{[g]} - \vec{d}_2^{[g]}$.) Figure 14 illustrates the results of a convergence and consistency test in which the (-1, -2) transmitted order's efficiency is plotted versus the order truncation limit **m_max** for each of the three illustrated stripe orientations, and for several values of the number N of stripes per semicircle used to define the pillar geometry. N has very little effect on the computation result, but there is a computed efficiency spread of about $2.6 \cdot 10^{-4}$ between the different stripe orientations (with **m_max** = 9).
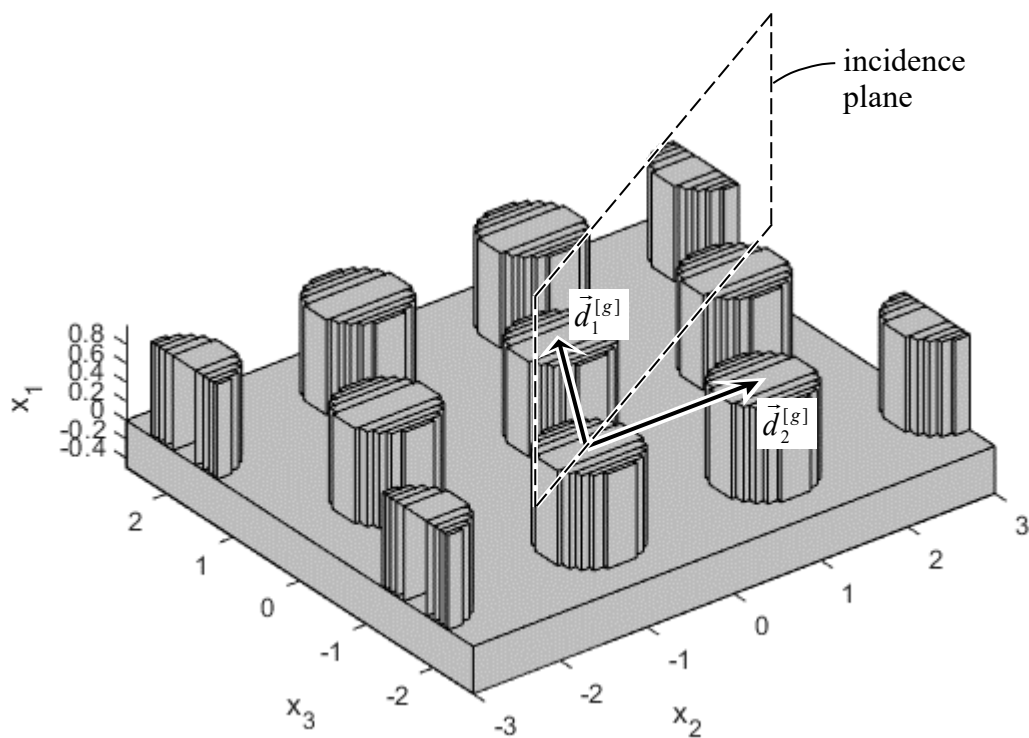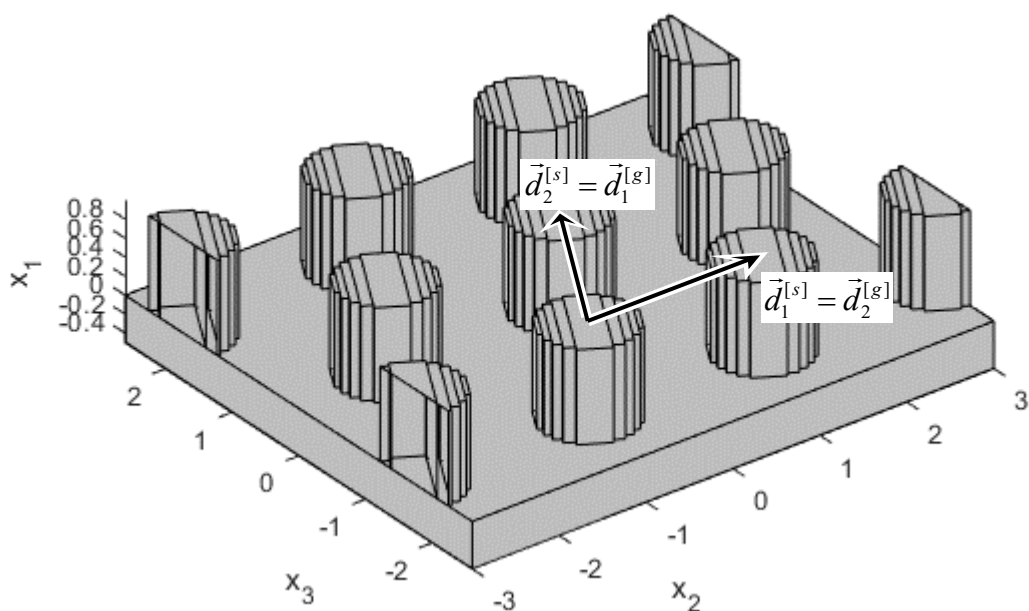
Figure 11. Pillar grating of demo 5.

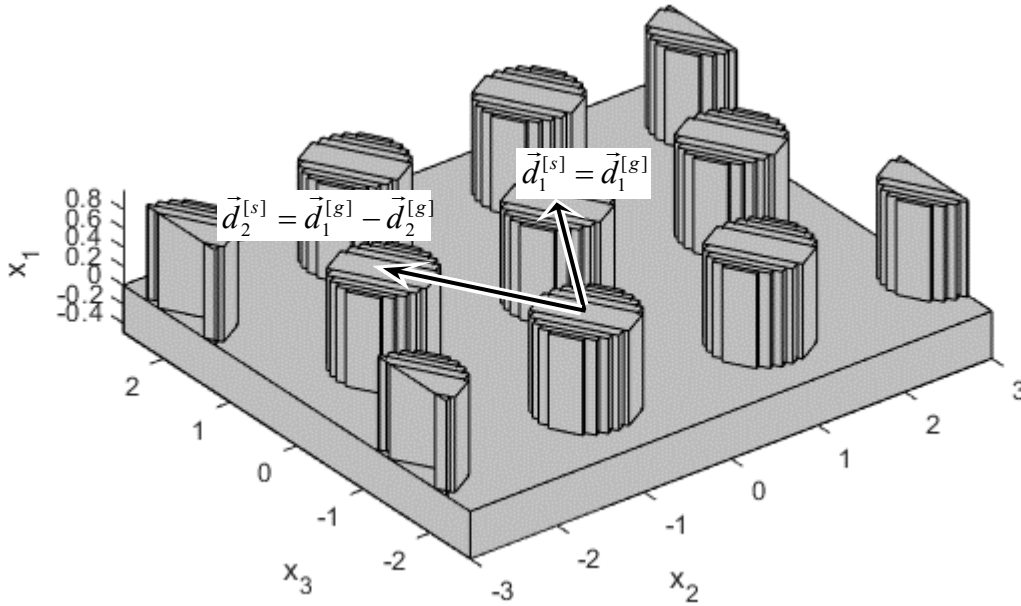Figure 12. Pillar grating of demo 5 with second stripe orientation.

Figure 13. Pillar grating of demo 5 with third stripe orientation.

In performing the above test, it is convenient to define the grating geometry in relation to the same grating periods ($\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$) and the same coordinate system for all three stripe orientations so that they can all use the same diffraction order indexing (i.e., index conversions such as Eq's. 18 and 19 are not required); and it is also convenient to use the same incident beam specification for all three configurations. This can be accomplished without having to use different pillar geometry specifications for the different stripe orientations – it is only necessary to change the values of several "harmonic indices" that define a stratum's periodicity in relation to that of the grating.

A stratum is characterized by two fundamental period vectors, $\vec{d}_1^{[s]}$ and $\vec{d}_2^{[s]}$, which are typically defined to match to the respective grating periods, $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$. However, they may be defined differently, either because the stratum exhibits stronger periodic symmetries than the overall grating, or – as in the case of demo 5 – because it is simply more convenient to use a different set of basis periods to define the stratum. The only limitation in defining the stratum periods is that $\vec{d}_1^{[g]}$ and $\vec{d}_2^{[g]}$ must be linear combinations of $\vec{d}_1^{[s]}$ and $\vec{d}_2^{[s]}$ with integer linear coefficients, i.e.,

$$\vec{d}_1^{[g]} = \vec{d}_1^{[s]} h_{1,1} + \vec{d}_2^{[s]} h_{2,1} \tag{26}$$

$$\vec{d}_2^{[g]} = \vec{d}_1^{[s]} h_{1,2} + \vec{d}_2^{[s]} h_{2,2} \tag{27}$$

where the h's ("harmonic indices") are integers (GD-Calc.pdf, Eq. 3.20). In other words, the stratum's basis spatial frequencies $\vec{f}_1^{[s]}$ and $\vec{f}_2^{[s]}$ must be harmonics of the grating frequencies $\vec{f}_1^{[g]}$ and $\vec{f}_2^{[g]}$,

$$\vec{f}_1^{[s]} = h_{1,1}\, \vec{f}_1^{[g]} + h_{1,2}\, \vec{f}_2^{[g]} \tag{28}$$

$$\vec{f}_2^{[s]} = h_{2,1}\, \vec{f}_1^{[g]} + h_{2,2}\, \vec{f}_2^{[g]} \tag{29}$$
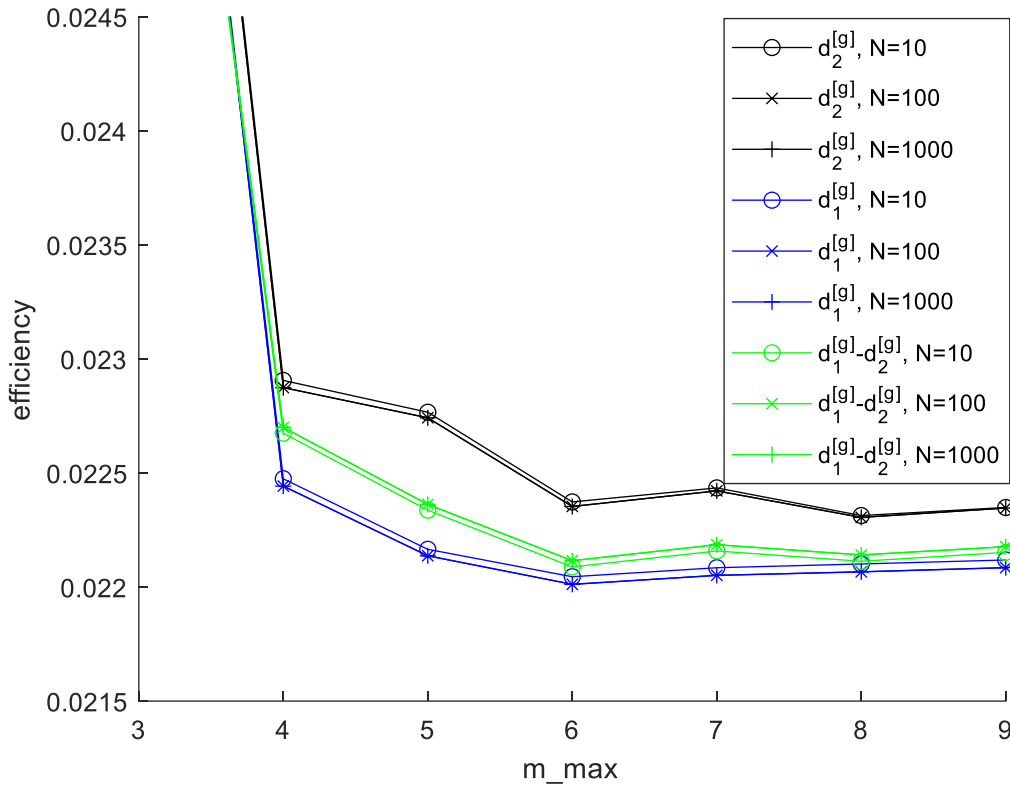
(GD-Calc.pdf, Eq. 3.33).

Figure 14. Convergence and consistency test for the pillar grating, (-1, -2) transmitted order, incident TM polarization. The legend indicates which vector is parallel to the stripes ($\vec{d}_2^{[g]}$, $\vec{d}_1^{[g]}$, or $\vec{d}_1^{[g]} - \vec{d}_2^{[g]}$). N is the number of stripes per semicircle used to define the pillar geometry. Hexagonal order truncation is used: $|m_1| \leq \mathtt{m\_max}$, $|m_2| \leq \mathtt{m\_max}$, $|m_2 - m_1| \leq \mathtt{m\_max}$.

A stratum's geometry is specified in relation to its period vectors (as illustrated in GD-Calc.pdf, Figure 3), with the stratum stripes parallel to $\vec{d}_2^{[s]}$. A linear coordinate transformation can be implicitly applied to the stratum geometry by changing the definitions of $\vec{d}_1^{[s]}$ and $\vec{d}_2^{[s]}$ (via the harmonic indices), and the definition can also be used

to determine the stripe orientation. In the Figure 11 configuration, the stratum periods are simply equated to the respective grating periods: $\vec{d}_1^{[s]} = \vec{d}_1^{[g]}$, $\vec{d}_2^{[s]} = \vec{d}_2^{[g]}$, so the stripes are parallel to $\vec{d}_2^{[g]}$. In Figure 12 the harmonic indices are redefined to swap $\vec{d}_1^{[s]}$ and $\vec{d}_2^{[s]}$, i.e. $\vec{d}_1^{[s]} = \vec{d}_2^{[g]}$, $\vec{d}_2^{[s]} = \vec{d}_1^{[g]}$, so the stripes are parallel to $\vec{d}_1^{[g]}$. In Figure 13, the harmonic indices are set to define $\vec{d}_1^{[s]} = \vec{d}_1^{[g]}$, $\vec{d}_2^{[s]} = \vec{d}_1^{[g]} - \vec{d}_2^{[g]}$, so the stripes are parallel to $\vec{d}_1^{[g]} - \vec{d}_2^{[g]}$. (The "**alt_stripe**" selector in gdc_demo5.m controls the branch code that defines the harmonic indices.)

Demo 6 and 7: Biperiodic grating comprising a skewed metal grid

Figures 15-18 illustrate the structure modeled by demo 6 and 7, which is a metal grid on a dielectric substrate. (For clarity of illustration the number of stripes is reduced in Figures 15 and 16.) Demo 6 models the structure using a unit cell designated as cell A, which is defined by the period vectors $\vec{d}_1^{[g,A]}$ and $\vec{d}_2^{[g,A]}$. Two stripe orientations are used, one with the grating stripes parallel to $\vec{d}_2^{[g,A]}$ (Figure 15) or $\vec{d}_1^{[g,A]}$ (Figure 16). Demo 7 uses a different unit cell, designated as cell B, whose defining period vectors $\vec{d}_1^{[g,B]}$ and $\vec{d}_2^{[g,B]}$ are

$$\vec{d}_1^{[g,B]} = \vec{d}_1^{[g,A]} - \vec{d}_2^{[g,A]} \tag{30}$$

$$\vec{d}_2^{[g,B]} = \vec{d}_1^{[g,A]} + \vec{d}_2^{[g,A]} \tag{31}$$

Again, two stripe orientations are used, one parallel to $\vec{d}_2^{[g,B]}$ (Figure 17) and the other parallel to $\vec{d}_1^{[g,B]}$ (Figure 18). (The "**alt_stripe**" selector in gdc_demo6.m and gdc_demo7.m controls the stripe orientation.)

Figure 19 illustrates a convergence and consistency test in which the efficiency of the (0, -1) transmitted order (relative to unit cell A) is plotted versus the order truncation limit for the four configurations of Figures 15-18, and for several values of the number N of stripes per grid hole. (Rectangular order truncation is used: $|m_1| \leq$ **m_max**, $|m_2| \leq$ **m_max** .) Figure 20 illustrates a similar test for the (0, 0) reflected order.
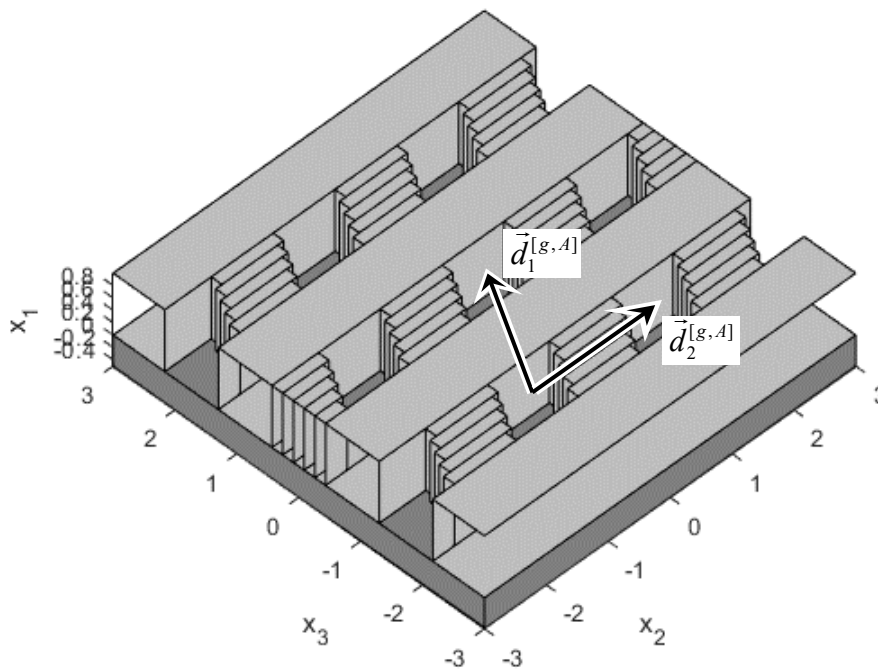
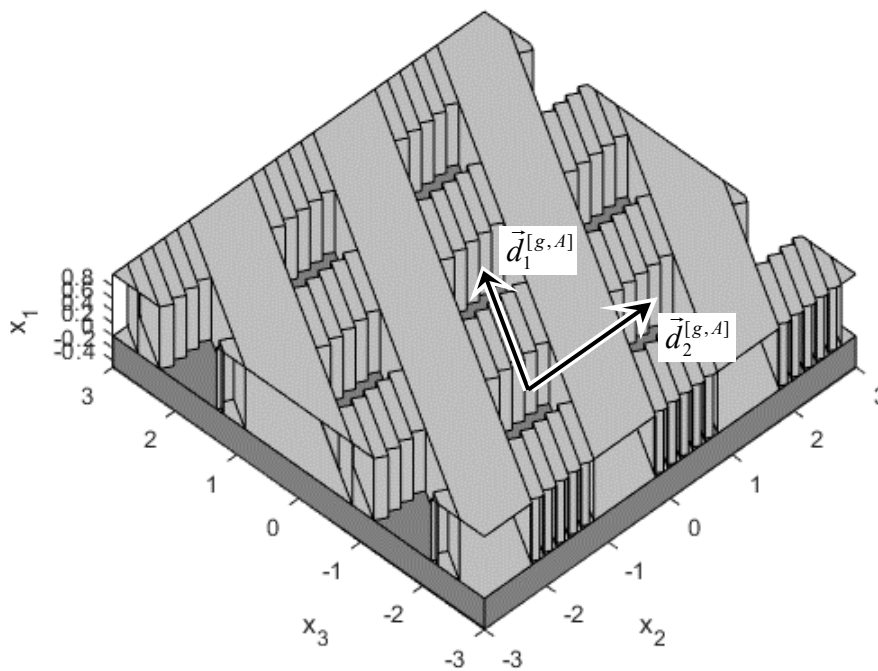Figure 15. Metal grid grating with unit cell A (demo 6), first stripe orientation.

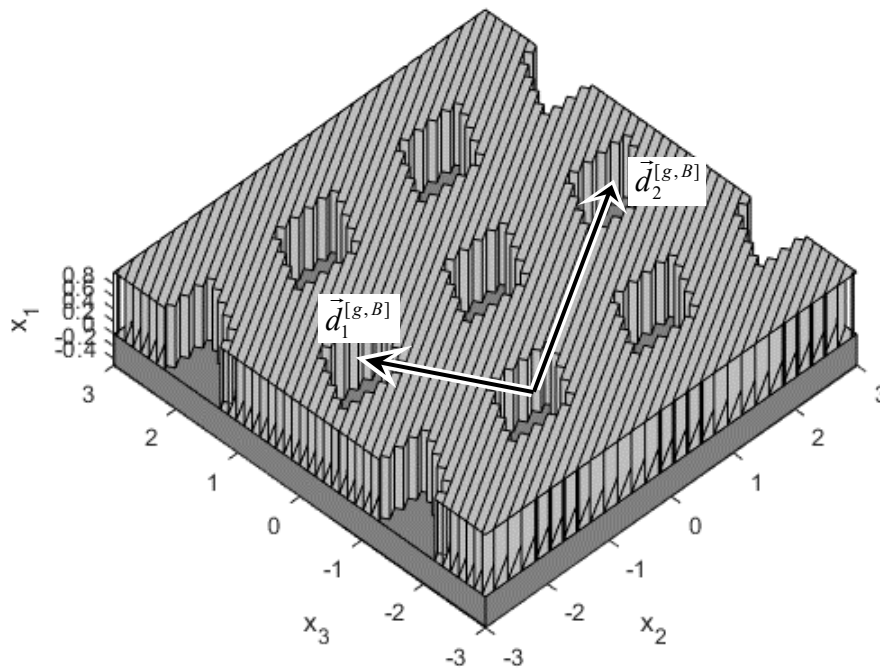Figure 16. Metal grid grating with unit cell A (demo 6), second stripe orientation.

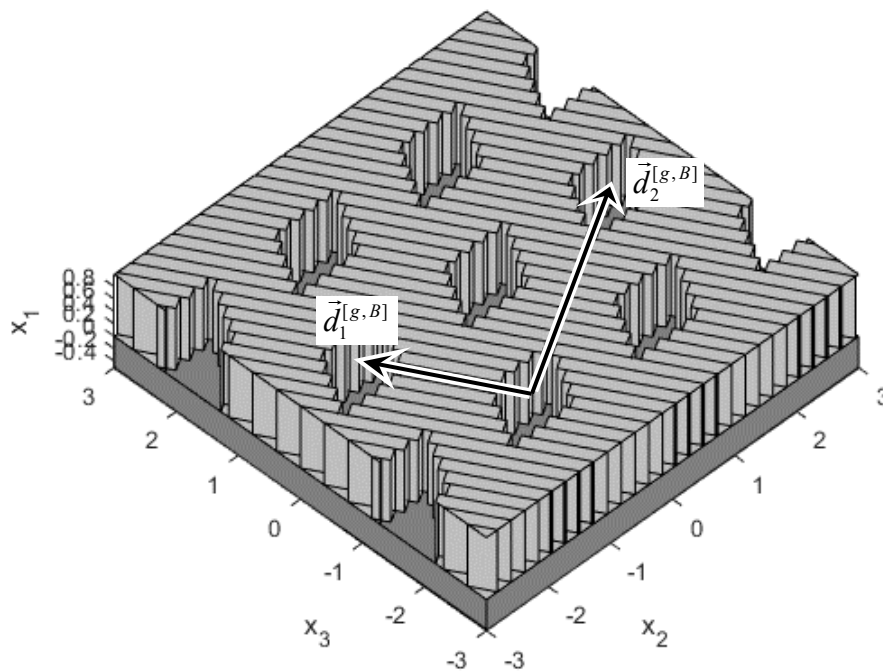Figure 17. Metal grid grating with unit cell B (demo 7), first stripe orientation.

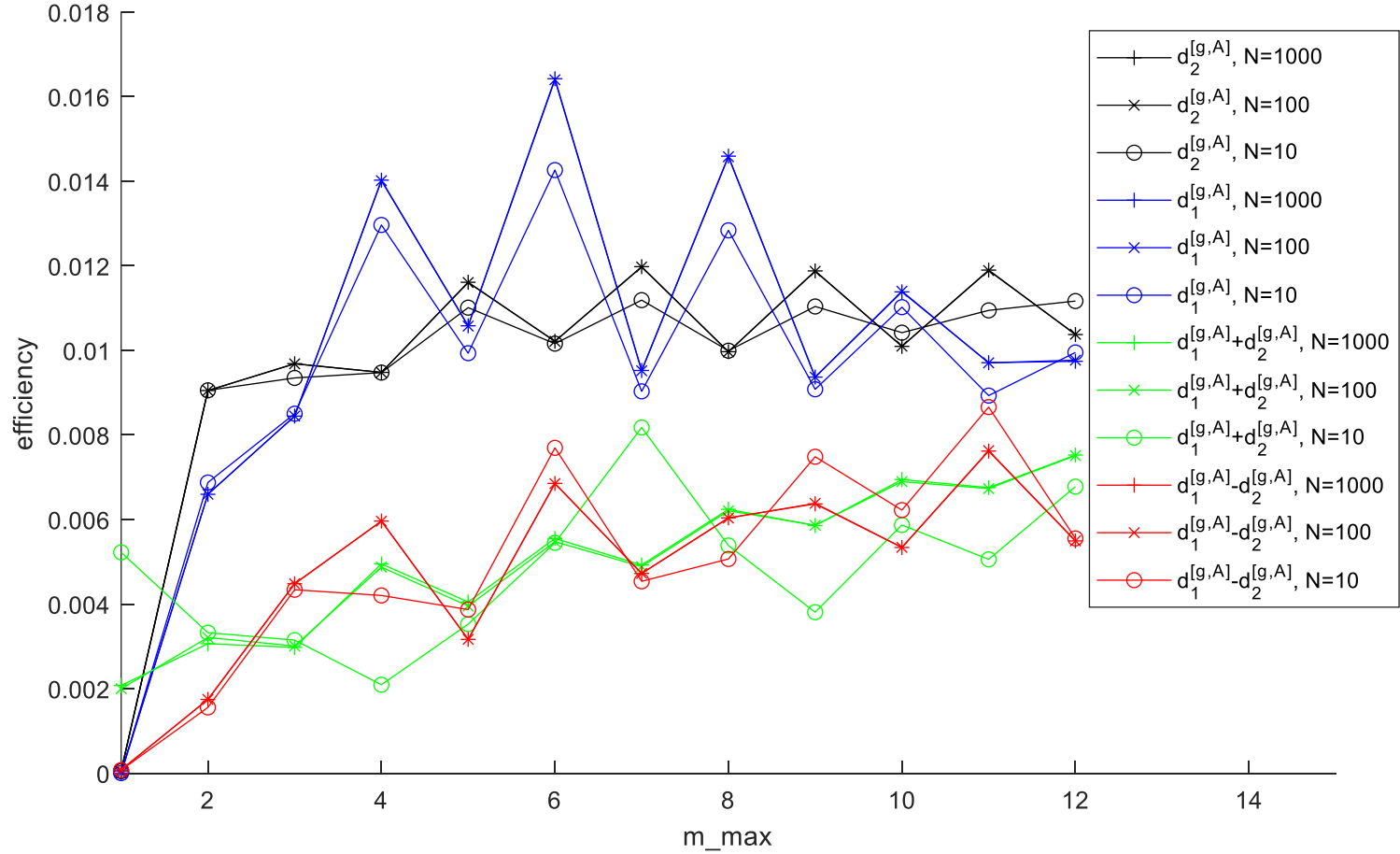Figure 18. Metal grid grating with unit cell B (demo 7), second stripe orientation.

Figure 19. Convergence and consistency test for metal grid grating of demo 6 and demo 7, (0, -1) transmitted order (relative to unit cell A), incident TM polarization. The legend indicates which vector is parallel to the stripes ($\vec{d}_2^{[g,A]}$, $\vec{d}_1^{[g,A]}$, $\vec{d}_1^{[g,A]} + \vec{d}_2^{[g,A]}$, or $\vec{d}_1^{[g,A]} - \vec{d}_2^{[g,A]}$). N is the number of stripes covering each hole.
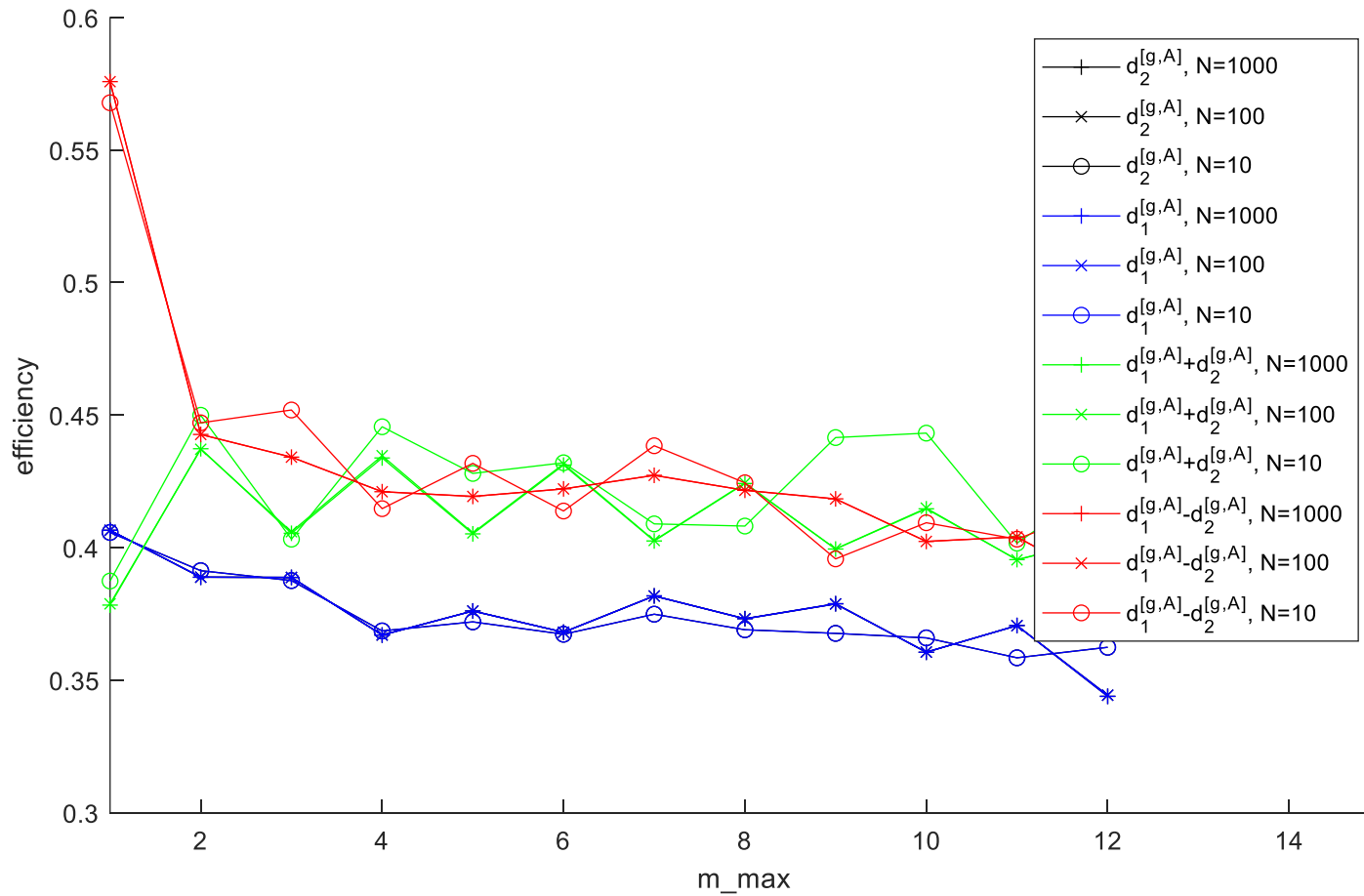
Figure 20. Same as Figure 19, but for the (0, 0) reflected order.

Demo 8: Biperiodic grating comprising a square metal grid

Figure 21 illustrates the grating structure of demo 8, which is a free-standing, square metal grid illuminated at normal incidence. The x_2 and x_3 coordinate axes are aligned to the grid, and the grating model's stripes are parallel to $\hat{e}_3$. Figure 22 illustrates a convergence and consistency test for this grating, using one of the first reflected diffraction orders.
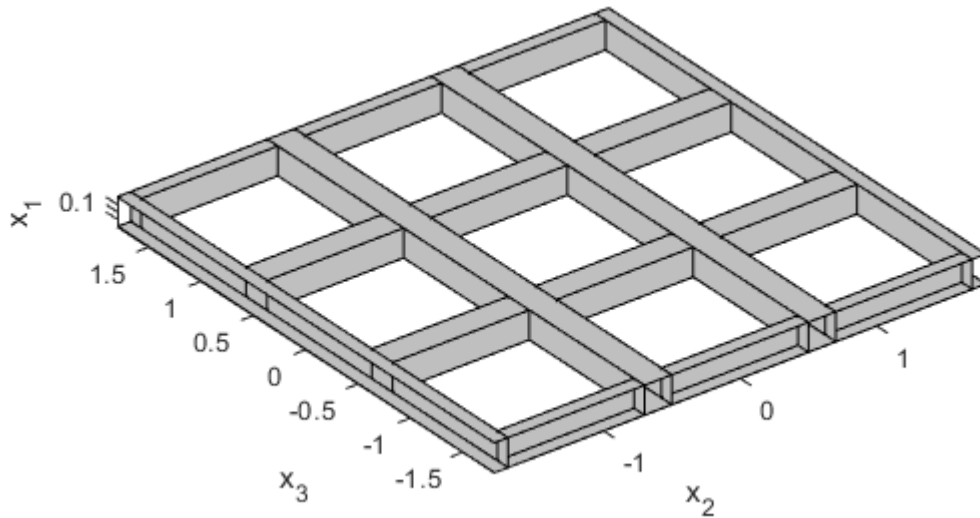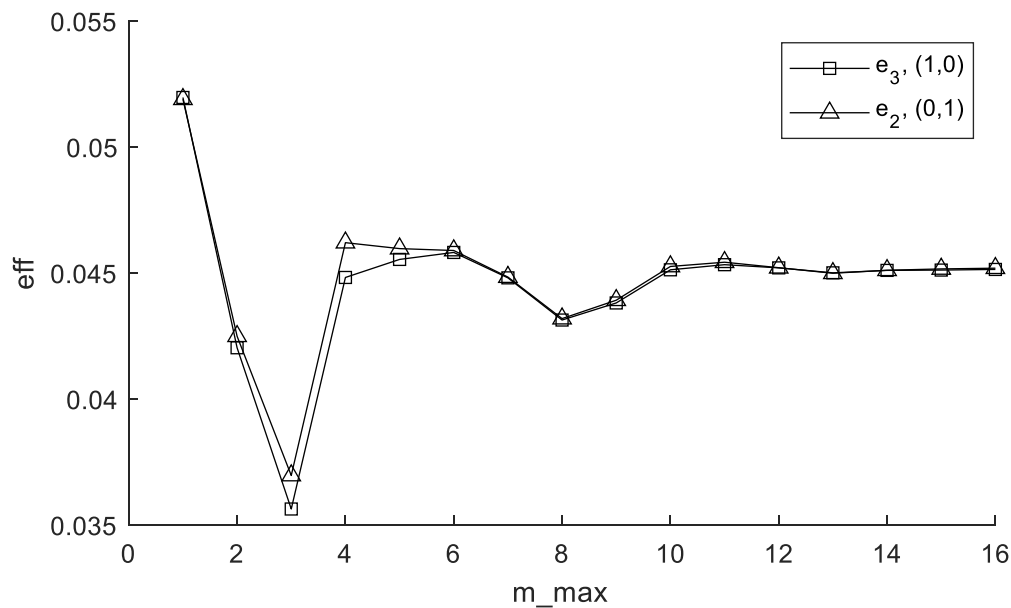

Figure 21. Square metal grid grating of demo 8.

Figure 22. Convergence and consistency test for demo 8. The legend indicates the incident polarization direction ($\hat{e}_3$ or $\hat{e}_2$) and reflected diffraction order.

Demo 9: Alignment sensor

This demo models an alignment sensor, which senses small lateral displacements between two proximate grating elements by comparing the intensities of the $+1^{st}$ and -1st-order diffracted beams generated by the gratings. Figure 23 illustrates the grating configuration. Normally-incident illumination passes through an upper transmission grating, reflects off of a lower reflection grating, and passes again through the upper grating, generating $+1^{st}$ and -1st diffraction orders whose relative intensity depends on the gratings' lateral displacement. The upper grating's superstrate, which is modeled as a semi-infinite dielectric medium, is not shown in Figure 23. Δx_2 is the lateral displacement between the upper and lower gratings' center lines. A plot of the computed alignment signals as a function of Δx_2 is illustrated in Figure 24.
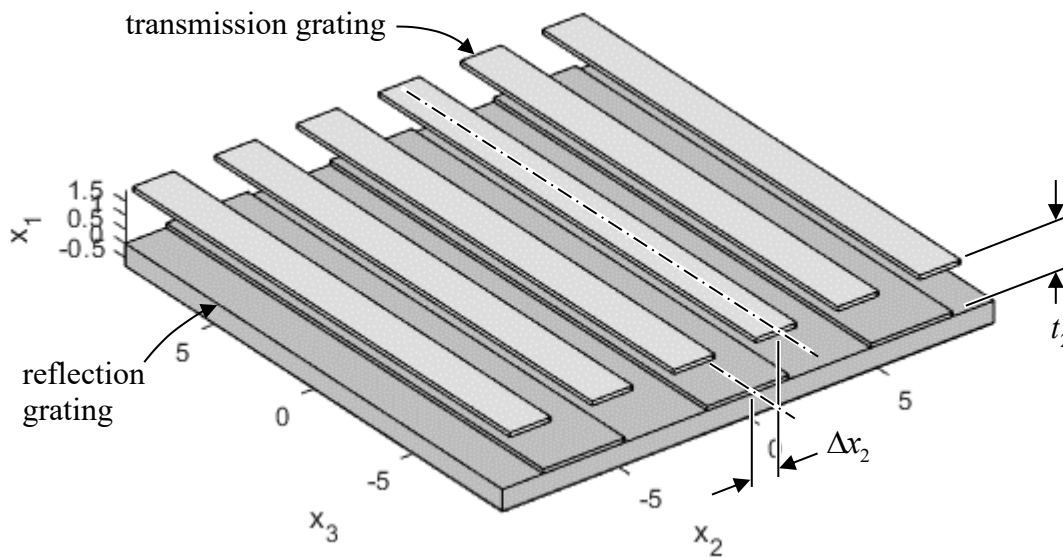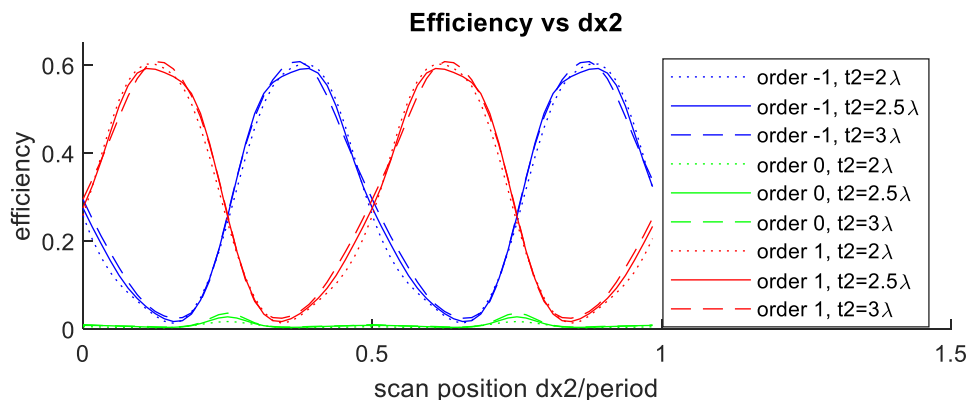

Figure 23. Alignment sensor.

Figure 24. Alignment sensor's efficiency in diffraction orders ± 1 and 0 (reflected, with incident field unpolarized) as a function of the lateral displacement Δx_2 (ratioed to the grating period), for three air spaces t_2 .

The GD-Calc interface defines a special stratum type that represents a coordinate break, which is used to define the lateral displacement of the top grating element in demo 9. The coordinate break is abstractly classified as a "stratum" of zero thickness, in the sense that it is associated with a lateral plane at a particular x_1 height in the grating, and it has the effect of applying a specified lateral (x_2, x_3) translational shift to all strata above the coordinate break.

The system illustrated in Figure 23 is represented in the demo script as a multilayer grating structure comprising four strata: (1) the reflection grating stratum, (2) the air gap of thickness t_2, (3) the coordinate break with translational shift Δx_2, and (4) the transmission grating stratum. The top grating stratum has a period half that of the bottom grating stratum, so it is specified with a harmonic index of 2, i.e., the top stratum's fundamental spatial frequency is twice that of the overall grating structure. (Only two harmonic indices, $h_{1,1}$ and $h_{1,2}$, are defined for a uniperiodic stratum – in this case $h_{1,1} = 2$ and $h_{1,2} = 0$ for the top stratum – because the stratum is characterized by a single frequency vector $\vec{f}_1^{[s]}$; cf. Eq. 28.)

Demo 9 illustrates the parameterization capabilities of GD-Calc. Quantities that are identified as "parameters" in the interface specification (see gdc.m comment header) can, in general, be numeric arrays of any dimensionality. Parameters must all be size-matched, except that a parameter's singleton dimensions are implicitly repmat-expanded to match parameter dimensions. (See the MATLAB help documentation on Array vs. Matrix Operations and Compatible Array Sizes for Basic Operations.) Typically, each parameter dimension is associated with a scalar variable. For example, in demo 9 the air space t_2 is defined as a size-[1, 3] row vector (`t2 = lambda*[2,2.5,3]`) and the lateral shift Δx_2 is defined as a size-[64,1] column vector (`dx2 = period*(0:63).'/64`). In the GD-Calc internal calculations, all numeric quantities that depend on t_2, but not Δx_2, will be size-[1, 3] row vectors; all that depend on Δx_2, but not t_2, will be size-[64,1] column vectors; and any that depend on both t_2 and Δx_2 will be size-[64, 3] matrices. In particular, the diffraction efficiency arrays returned by `gdc_eff` will generally be size-[64, 3], as implied by the description of the efficiency outputs as "parameters" in gdc_eff.m. (Zero efficiency values might be represented as scalar zeros, but the values can be expanded to size-[64, 3] matrices via repmat-expansion of singleton dimensions.)

Parameterization can, in some instances, considerably improve computation runtime, but it can also increase memory requirements. Therefore, parameterization should be used judiciously on complex, memory-intensive problems, such as biperiodic gratings, to get the most performance benefit within the limitation of memory resources. Some understanding of how GD-Calc's internal algorithms work can provide insights into how to use parameterization efficiently.

The GD-Calc calculations are primarily based on two operations: first, calculating an "S matrix" (scattering matrix) for each individual stratum in the grating; and then combining the S matrices for the strata by means of a "stacking" operation, which is applied from bottom to top to determine a cumulative S matrix for the entire grating. In demo 9 there are $3 \times 64 = 192$ parameter combinations in total (3 air spacings and 64 lateral displacements), so if the t_2 and Δx_2 parameters were iterated outside gdc.m, all of these S matrix computations would have to be repeated 192 times. But if vectorized parameters are used the S matrix for the grating stratum will only be calculated once (because it has no dependence on t_2 or Δx_2), the S matrix for the air space (which depends only on t_2) will be calculated only 3 times, and the S matrix for the coordinate break (which depends only on Δx_2) will only be calculated 64 times. Furthermore, the air space stacking operation will only be calculated 3 times, and only the stacking operations above the air space must be repeated for all 192 parameter combinations.

Generally, parameter vectorization can be most effective when applied to parameters that only affect the uppermost grating strata, because this will avoid redundant stacking operations for lower strata. Thus, for complex, resource-intensive applications a good strategy would be to first try vectorizing parameters for the top strata, and then try vectorizing parameters associated with lower strata.

The default bottom-to-top stacking order can be changed by encapsulating some strata in a replication module (stratum type 4 with replication count = 1). The module's composite S matrix will be calculated (via stacking S matrices of encapsulated strata) before stacking with lower strata. However, there is a factor-of-two runtime efficiency loss in the module's internal stacking operations because they must account for incident fields entering the module from both above and below, whereas the grating as a whole has no incident fields in the substrate.

Aside from avoiding duplicate operations, there is another way that parameterization can improve performance. The GD-Calc code implementation uses paged matrix operations (**pagemtimes**, **pagemldivide**, etc.), which parallelize matrix operations on multicore and multiprocessor computers. The runtime benefit of these operations depends on the number of diffraction orders (which is controlled by **m_max** in the demo scripts) and the number of processors and cores.

The plotting routine gdc_plot.m is designed to work with parameterized grating structures. The second **gdc_plot** argument is a multidimensional parameter index, which specifies which parameter combination is to be plotted. The demo 9 code illustrates how the plotting function can be used to generate a simple animation of Figure 23 showing a parameter scan of Δx_2.

Demo10: Holographic volume grating

Demo 10 models a holographic volume grating – either a transmission grating or a reflection grating; see Figure 25. The structure comprises a sinusoidal refraction index density variation, and the structure operates by distributed Bragg diffraction. The calculated diffraction efficiency, using both GD-Calc and the Kogelnik equations, is illustrated in Figure 26 for a transmission grating and in Figure 27 for a reflection grating.

The grating is approximated as a stack of thin, lamellar grating strata, all identical but with a lateral translation offset between strata. The lateral offset is represented by a coordinate break. Within each stratum the continuous refractive index variation is approximated by a piecewise constant function. (The number of strata and discretization steps are reduced in Figure 25 for the purpose of illustration.)
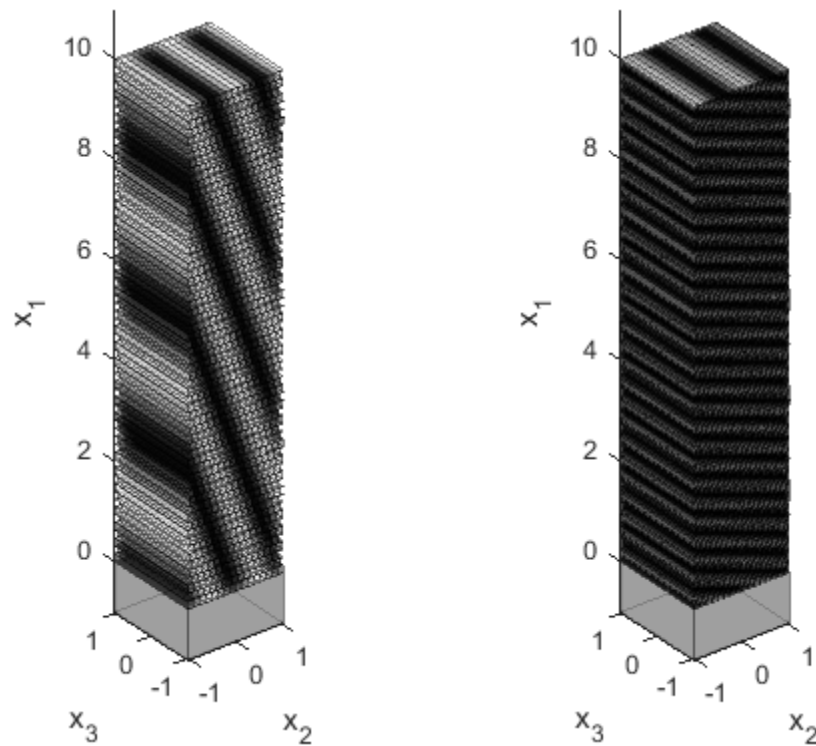

Figure 25. Holographic volume diffraction grating of demo 10 (left: transmission grating, right: reflection grating).
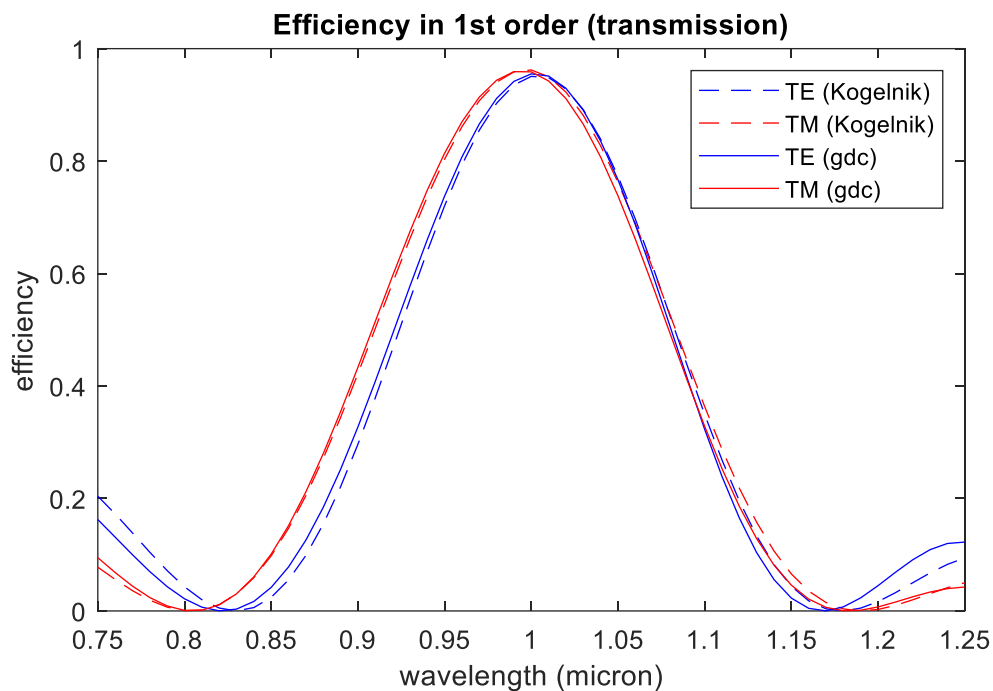
Figure 26. Volume transmission grating's Bragg diffraction efficiency in the 1st diffraction order for TE- and TM-polarized, normally-incident illumination.
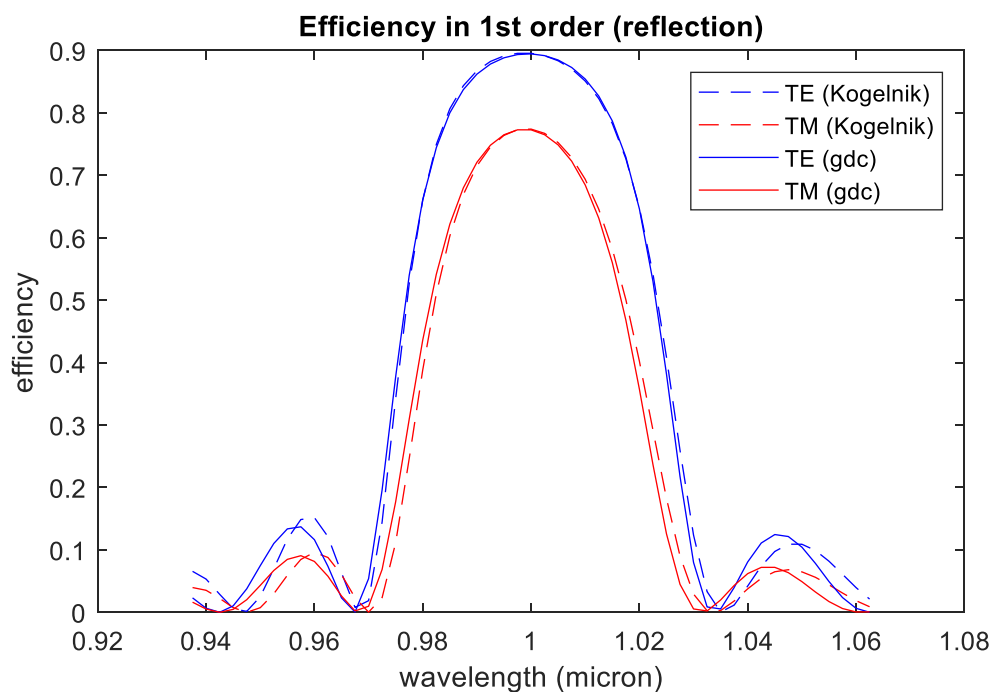

Figure 27. Volume reflection grating's Bragg diffraction efficiency in the 1st diffraction order for TE- and TM-polarized, normally-incident illumination.

Rather than explicitly defining separate lamellar strata and associated coordinate breaks, demo 10 uses a "replication module" to implicitly stack multiple copies of a structure comprising just one lamellar stratum and one coordinate break. The advantage of using a replication module is not just one of convenience: Whereas the computation runtime for S matrix stacking would generally be proportional to the number of strata, with a replication module the runtime is only proportional to the logarithm of the number of strata. Thus, a large number of very thin strata can be used without incurring a significant increase in runtime.

Demo 11: Crossed-line grating

The structure modeled by demo 11 (see Figure 28) is a photonic crystal comprising a stack of crossed, uniperiodic grating layers, each layer comprising square-section tungsten rods. (Refractive index data for tungsten is obtained from the file W.nk using the accessory function, read_nk.m.)

For this example, diagonal order truncation is used. Figure 29 illustrates convergence plots comparing the performance of rectangular truncation ($|m_1| \leq$ **m_max** and $|m_2| \leq$ **m_max**) and diagonal truncation ($|m_1| + |m_2| \leq$ **m_max**). The two methods exhibit similar convergence performance, but computation time is reduced by about a factor of 8 using diagonal truncation. Figure 30 illustrates the grating's computed zero-order reflectance and transmittance spectra for several values of the order truncation limit **m_max**.

The demo 11 code only computes the efficiency at one wavelength. Multiple-wavelength calculations (e.g., for Figure 30) can be done by either vectorizing wavelength or by iterating wavelength in an outer loop, but due to the high number of diffraction orders an external loop would probably be more practical and efficient.

The grating structure is represented as a replication module comprising the following three stratum elements (from bottom to top): (1) a uniperiodic stratum with the grating lines in the x_3 direction, (2) a uniperiodic stratum with the grating lines in the x_2 direction, and (3) a coordinate break, which applies a half-period translational shift in both the x_2 and x_3 directions. (The second uniperiodic stratum is constructed by simply swapping the harmonic indices of the first stratum.) The replication count (**rep_count**) is 4, giving a total of 8 uniperiodic grating strata. (The S matrix stacking time scales in proportion to log(**rep_count**), so much deeper structures could be easily analyzed.)
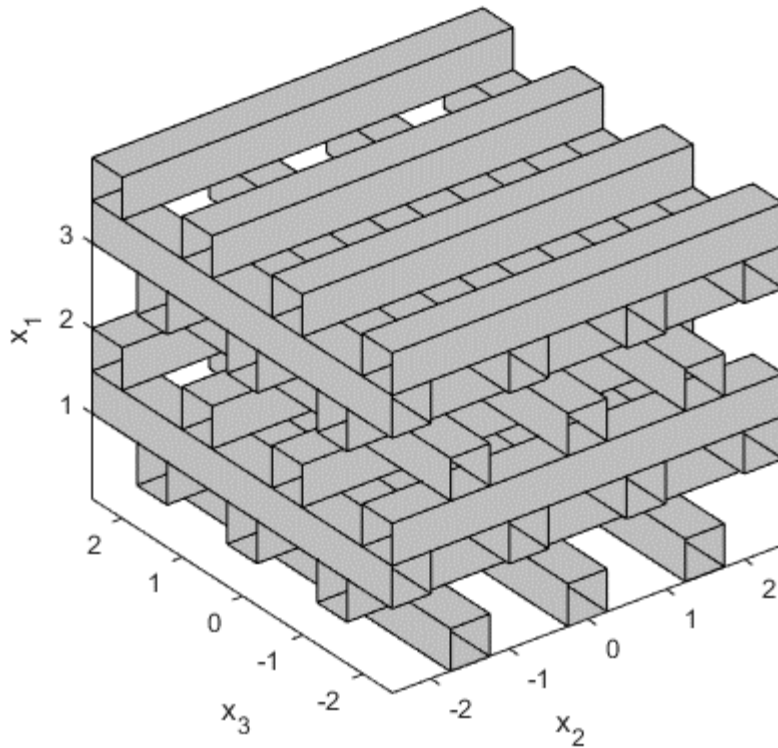
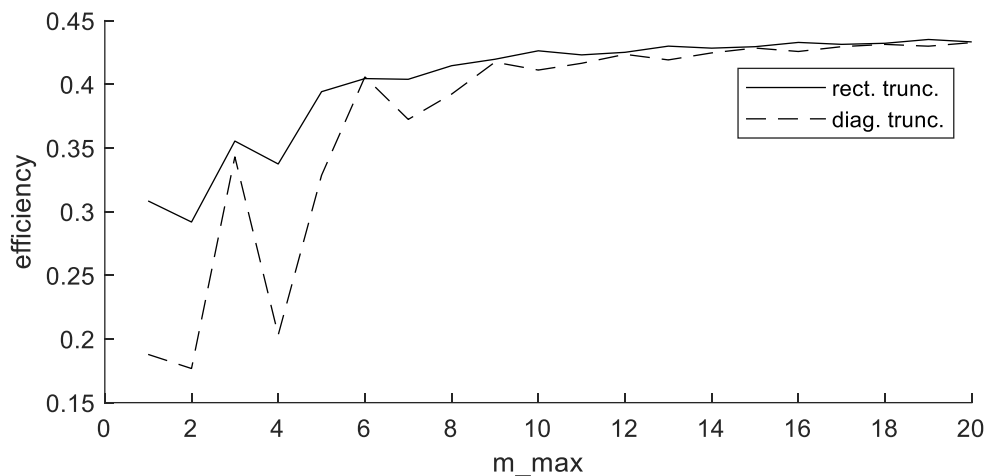Figure 28. Tungsten photonic crystal (demo 11).

Figure 29. Convergence test for tungsten photonic crystal (zero-order reflectance for unpolarized, normally-incident illumination at wavelength $1.825\,\mu\text{m}$). Two order truncation methods are compared: rectangular truncation ($|m_1| \leq$ **m_max** and $|m_2| \leq$ **m_max**) and diagonal truncation ($|m_1| + |m_2| \leq$ **m_max**).
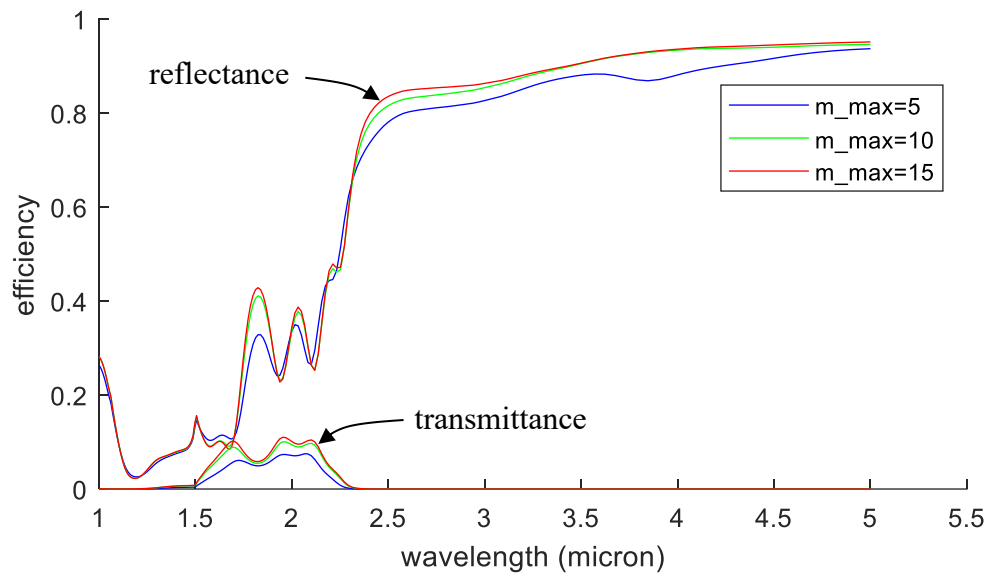
Figure 30. Computed zero-order reflectance and transmittance spectra for the tungsten photonic crystal, with unpolarized, normally-incident illumination, using diagonal order truncation.

Demo 12: Blazed, phase-Fresnel transmission grating

This demo models a blazed, phase-Fresnel transmission grating illustrated in cross-section in Figure 31. The grating is designed to achromatize a prism. Appendix B describes the design methodology.
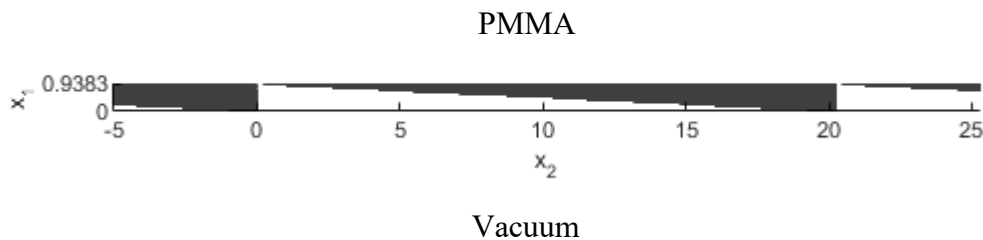

Figure 31. Blazed transmission grating.

The grating is a surface relief pattern in acrylic (PMMA), and the illumination transmits from inside the PMMA into vacuum (top to bottom in Figure 31). See Appendix B, Figure B1(b) for a more detailed view of the prism design geometry. The prism's exit surface transmission efficiency (for unpolarized incident light) is illustrated in Figure 32 without and with the grating compensator. There is a loss of about 5% due to surface reflections, and an additional 5% loss (at the blaze wavelength) from scattering

into extraneous diffraction orders. An antireflection coating would reduce the surface reflection loss, and an additional 2% efficiency gain could be achieved by increasing the facet wall angle from 90° to 116°.
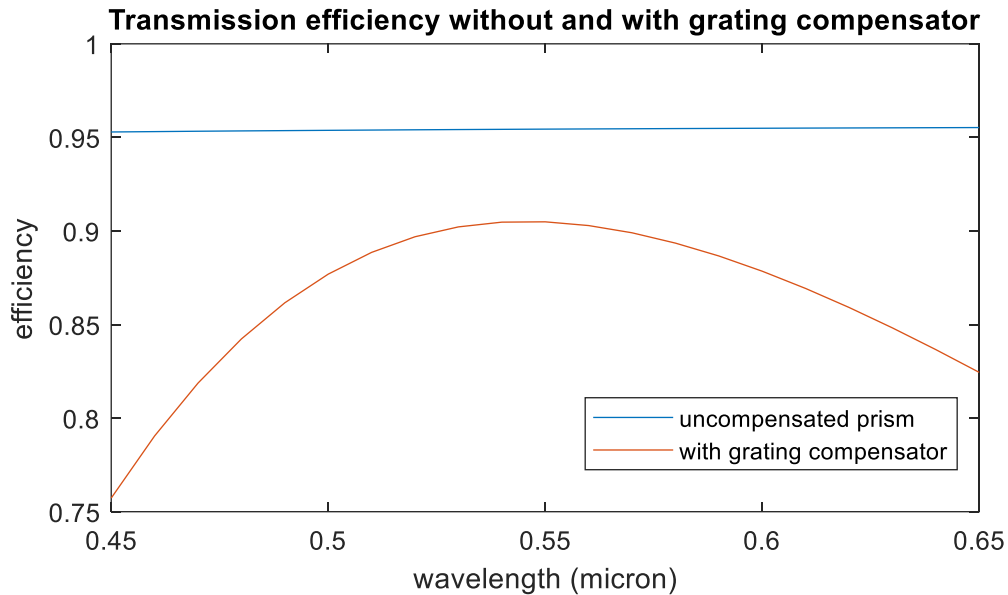

Figure 32. Transmission efficiency of prism exit surface without and with the achromatizing grating compensator.

Demo 13: Alignment sensor with internal field calculation

This demo repeats the alignment sensor example of demo 9 and adds internal field calculation. The animation of the grating motion is modified to also show one of the time-varying electromagnetic field components at the top of the air gap between the two gratings. Full-field calculation is enabled by inserting the following code line in the second stratum (air space) definition:

```
grating.stratum{2}.full_field = [];
```

gdc is called with the **grating** struct included in both the input and output argument lists, and the return value will have **full_field** initialized to a struct containing Fourier coefficients of all of the $\mathbf{E}$ and $\mathbf{H}$ field components, for TE and TM incident field polarizations. (See gdc.m comment header under "Advanced features – Full-field option" for details.) The Fourier coefficients for one of the field components ($\hat{e}_3 \cdot \mathbf{H}$, with a TM-polarized incident field) are converted to a spatial map via the function EH_map.m, and the spatial map (H_3 as a function of x_2) is plotted as a time-varying animation correlated with the grating scan. One frame of the animation is illustrated in Figure 33.
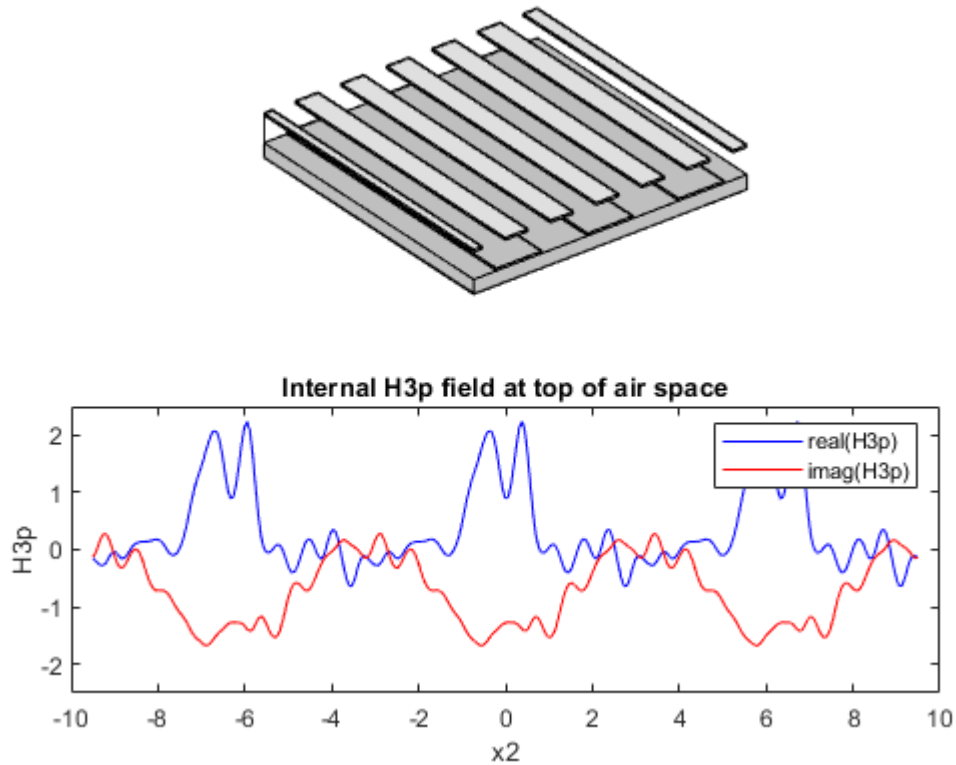
Figure 33. Demo 13 animation frame.

Demo 14: Biperiodic grating with internal field calculation

This demo calculates internal fields in a biperiodic grating comprising square gold posts on an SiO_2 substrate, with normally-incident radiation at wavelength 0.6328 micron polarized in the $\hat{e}_3$ direction. A horizontal (x_2, x_3) spatial map of each field component's magnitude ($|E_1|$, $|E_2|$, $|E_3|$, $|H_1|$, $|H_2|$, $|H_3|$) is generated at multiple height levels (x_1 positions) within the substrate, the grating, and the superstrate, and an animation is generated showing the total (incident plus diffracted) field variations with x_1.

The grating layer is split into ten strata to allow field sampling at multiple heights. Ten homogeneous strata are also added to the substrate, and similarly to the superstrate, to enable field sampling in the substrate and superstrate. Stratum **j** is initialized with the following assignment to activate field sampling:

```
grating.stratum{j}.full_field = [];
```

The **grating** struct returned by **gdc** includes **full_field** structs defining the **E** and **H** fields' Fourier coefficients within each stratum, at the top stratum boundary, and the Fourier coefficients are converted to spatial maps via the function EH_map.m. Figure

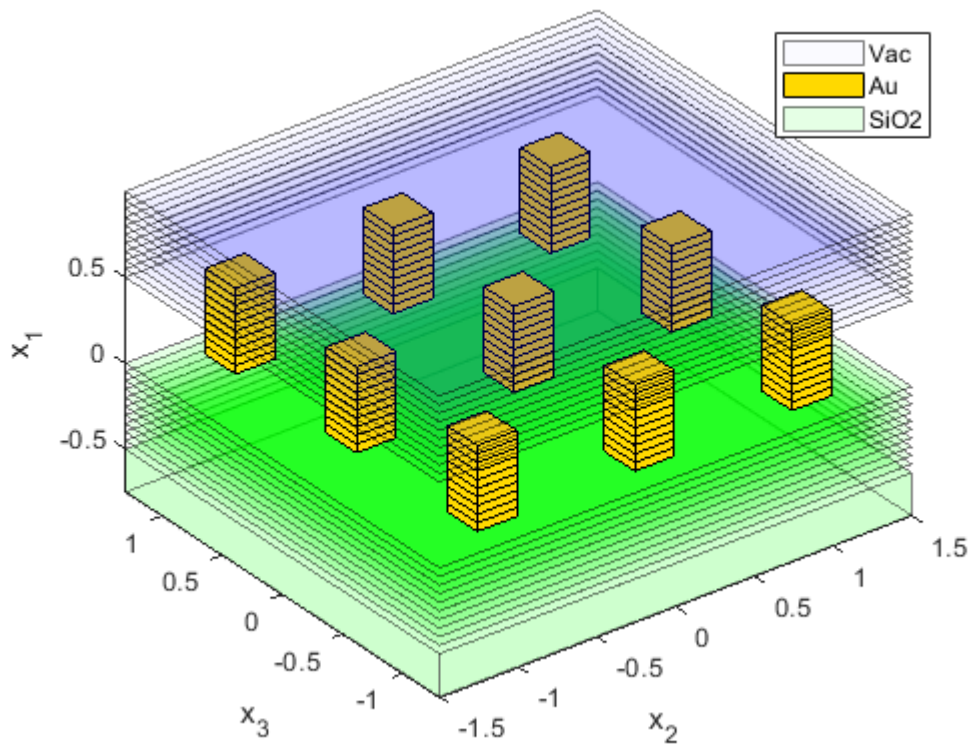34 shows the grating structure and the 30-level stratification, and Figure 35 shows one frame of the field animation.

Figure 34. Biperiodic gold grating structure for Demo 14, with stratification for field sampling.
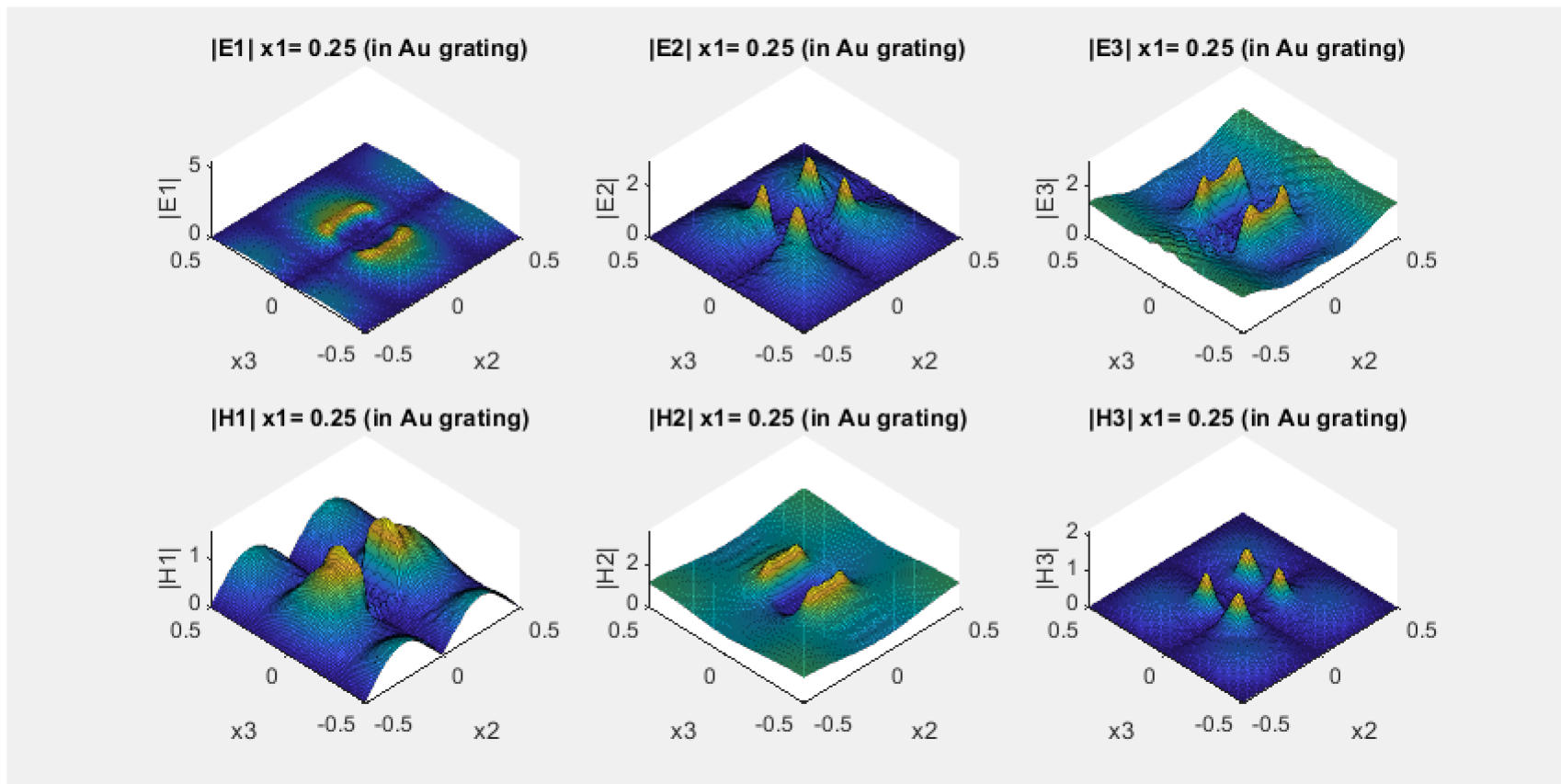
Figure 35. Demo 14 animation frame

Demo 15: Uniperiodic, sinusoidal grating with internal field calculation

This demo calculates and plots the electromagnetic field above a metal, sinusoidal relief grating, for TM polarization. The structure is approximated as a stack of 10 lamellar gratings, as illustrated in Figure 36, and each lamellar layer is further subdivided into 20 strata for internal field sampling. Figures 37(a)-(c) illustrate spatial maps of the H_3, E_1, and E_2 field magnitudes across the grating cross section (x_1, x_2 plane).
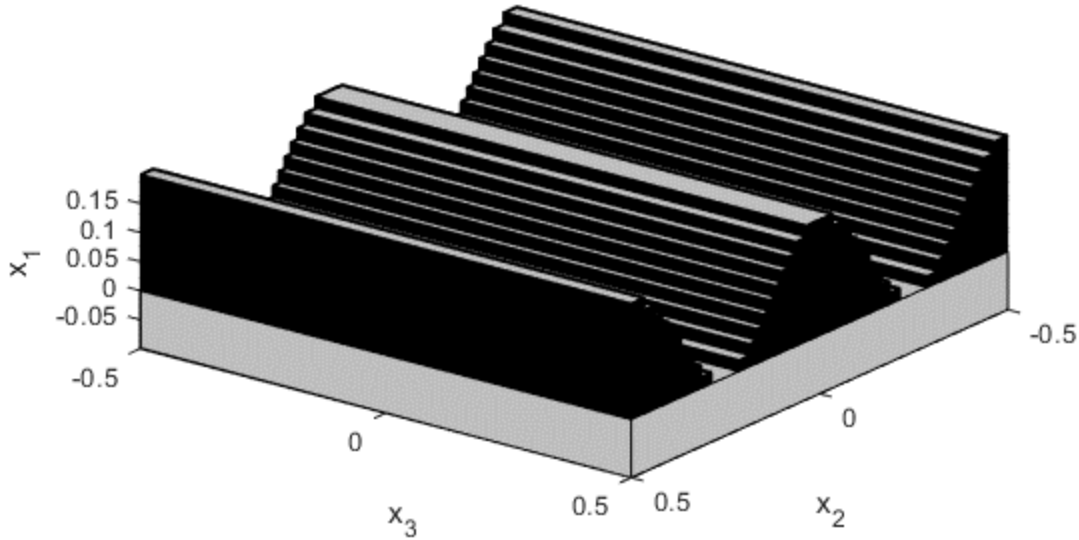
Figure 36. Sinusoidal relief grating with staircase approximation (10 steps) for demo 15
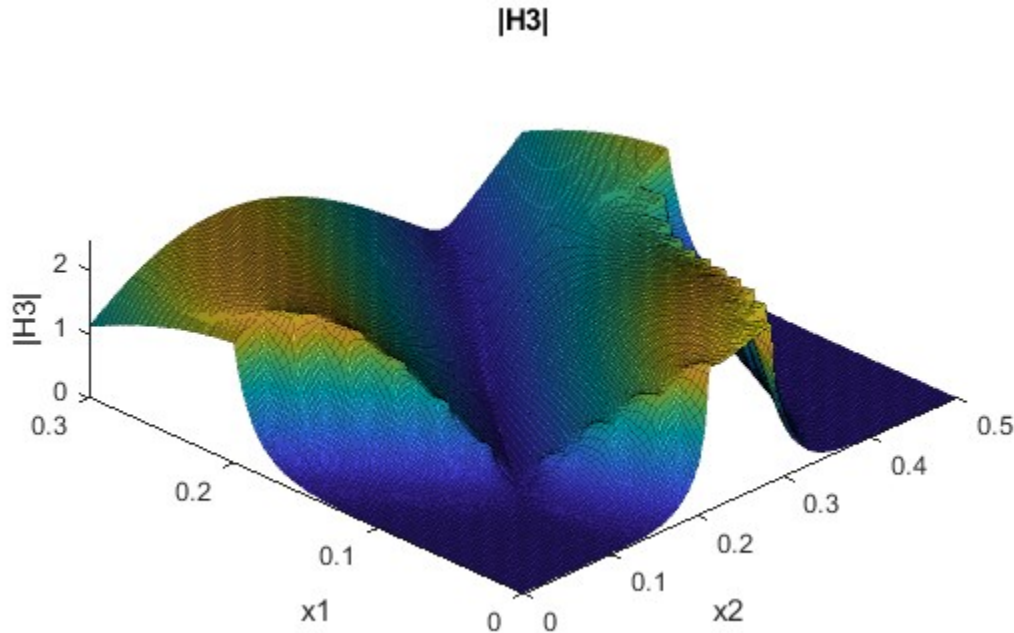
Figure 37(a). $|H_3|$ cross section through the Figure 36 grating

|E1|
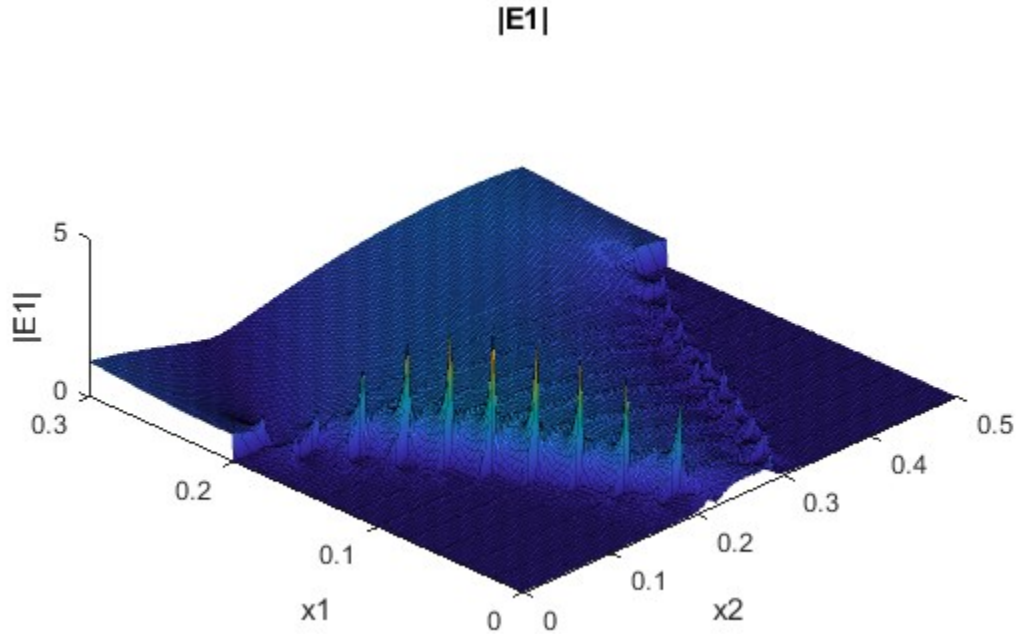
Figure 37(b). $|E_1|$ cross section through the Figure 36 grating

|E2|
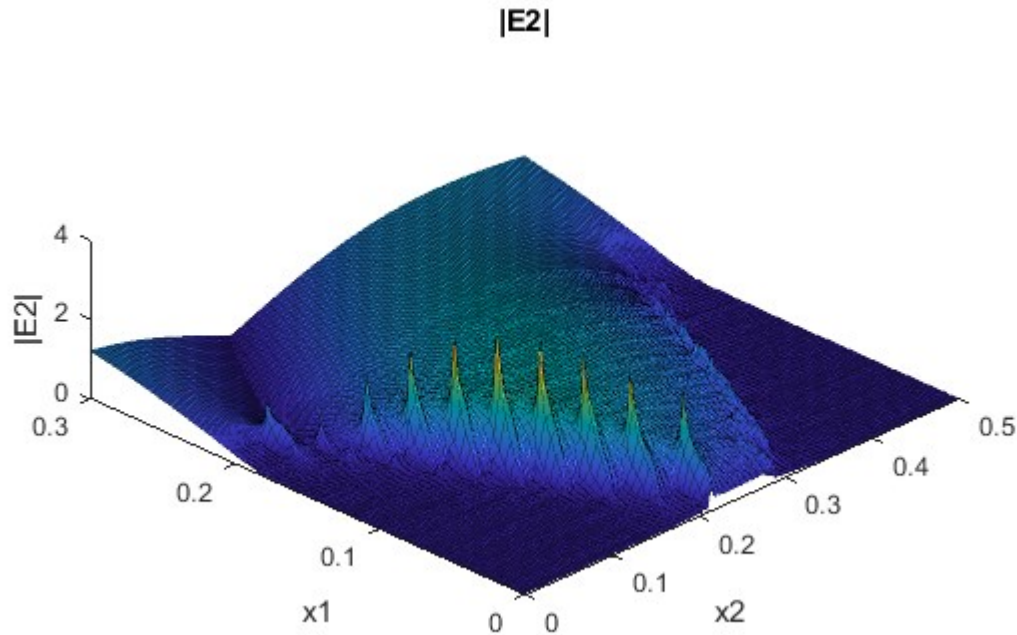
Figure 37(c). $|E_2|$ cross section through the Figure 36 grating

Figures 37(b) and 37(c) illustrate a limitation of the staircase approximation: A structure with actual physical steps would exhibit field singularities at the edges, which require many diffraction orders to resolve. As the number of steps is increased, the number of Fourier field order required to achieve convergence correspondingly increases and can become impractically high – if convergence can even be attained.

Demo 16: Coherent beam combiner

This demo illustrates a simple diffraction calculation with multiple incident beams. A dielectric transmission grating is designed to diffract a zero-order incident beam into two transmitted beam (diffraction orders 0 and 1) with equal intensity, Figure 38(a). Under coherent illumination by two incident beams (orders 0 and 1, Figure 38(b)) the power balance between the two transmitted beams depends on the incident beams' relative phase and the grating's translational position. Figure 39 illustrates the transmitted power modulation in the two beams as a function of the incident order 1 phase.
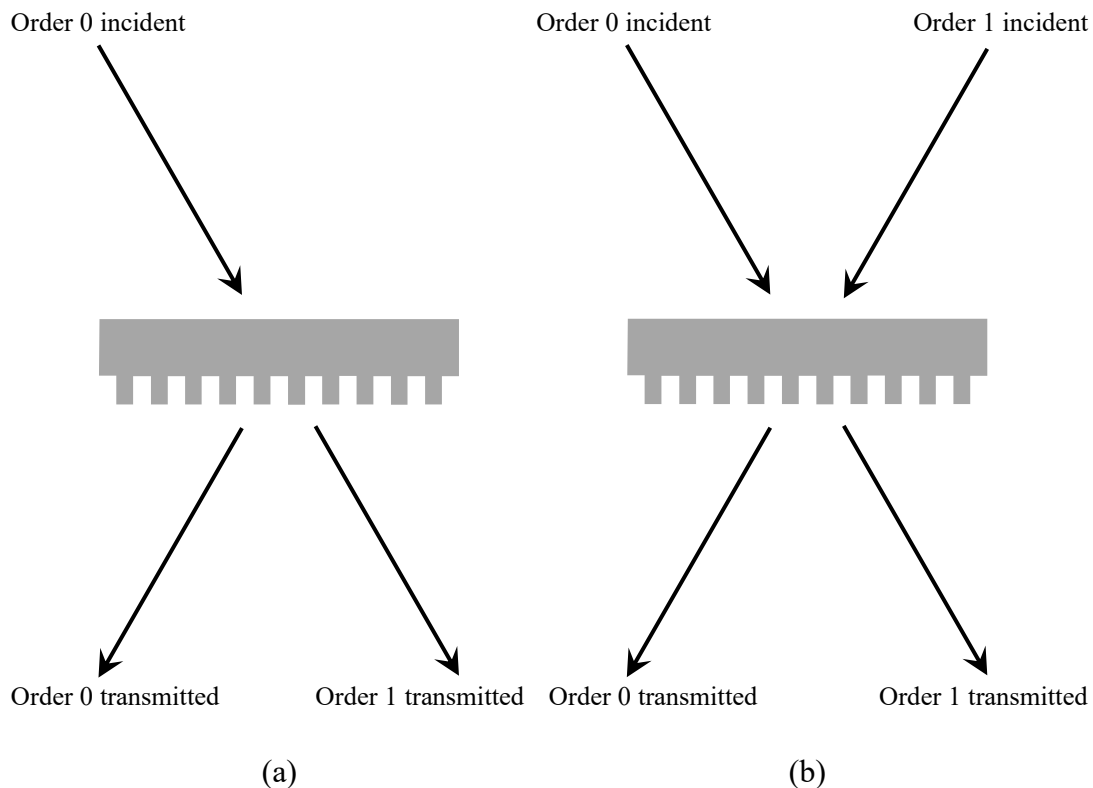
Figure 38. Transmission grating with one incident order (a) and with two incident orders (b).

gdc normally returns output data characterizing the grating for a single plane-wave incident beam. But the **gdc** internal calculations generate diffraction characteristics for all incident Fourier orders corresponding to the grating's spatial frequencies. (The incident orders' grating-tangential spatial frequencies match those of the diffraction orders.) An optional **gdc** input, **inc_order**, can be used to obtain diffraction characteristics for all, or a selected subset, of incident orders. Demo 16 uses this facility to acquire diffractive scattering coefficients for incident orders 0 and 1. The diffraction amplitudes from the two incident beams are summed to determine the coherently combined amplitudes.
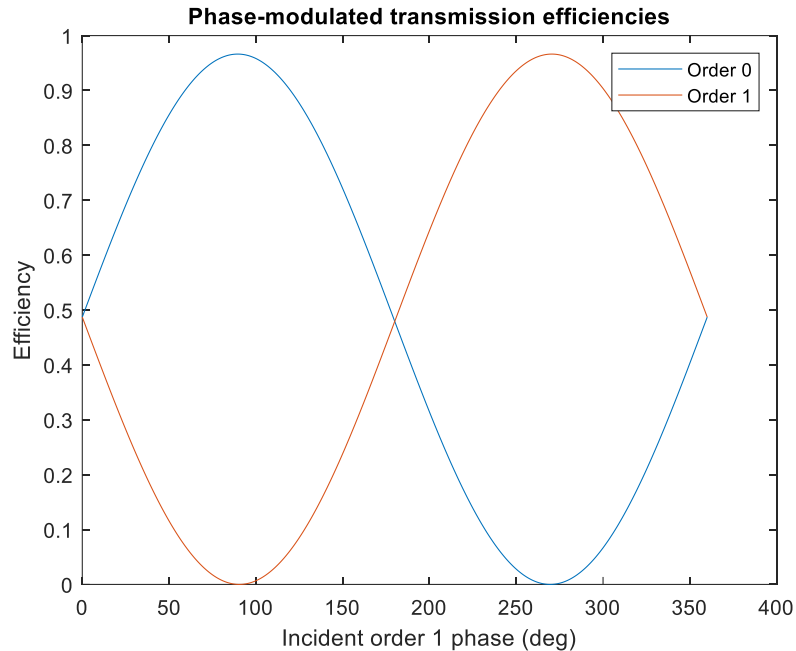
Figure 39. Coherent beam splitter's output beam modulation as a function of the incident order 1 phase.

Demo 17: Diffractive beam splitter with Gaussian beam

A limitation of the **inc_order** option illustrated in demo 16 is that it only models multiple plane-wave incident beams whose spatial frequencies differ by multiples of the grating's fundamental frequency. A general non-plane-wave incident beam such as a Gaussian beam can be approximated as a Fourier series of plane waves, but the frequency sampling is much denser than the grating frequencies. In this case, **inc_field.f2** and **inc_field.f3** in the **gdc** function's **inc_field** argument can be parameterized to parallelize diffraction calculations over the incident beam's Fourier components. As illustrated in Figure 40(a), the incident zero order comprises a cluster of plane waves spanning the beam's spatial frequency spectrum, and each diffracted order similarly comprises a cluster of plane waves. The incident waves can be assigned amplitudes to represent any spatial field map and polarization state.
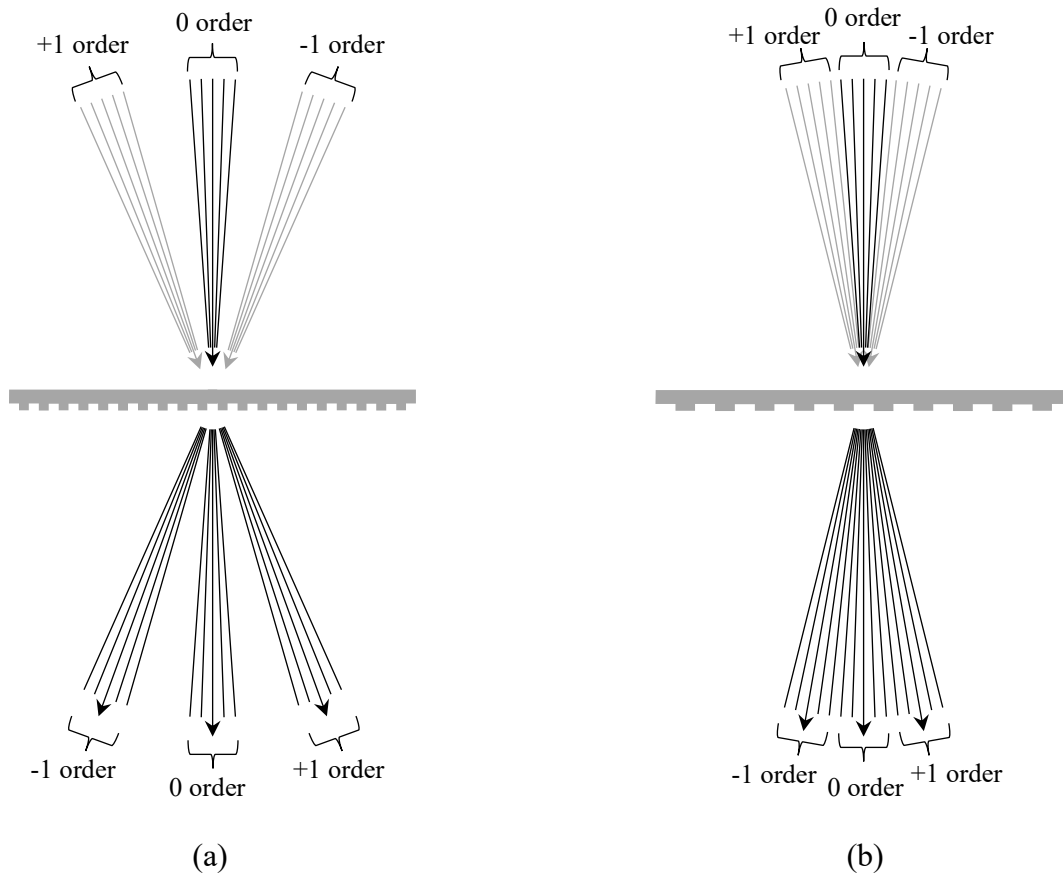
Figure 40. Multiple incident beams represented by **`inc_field.f2`** and **`f3`** parameterization, and by order separation (**`inc_order`**).

    If the **`inc_order`** option is used, the range of incident wave frequencies is extended to include the zero-order frequencies (**`f2`** and **`f3`**) added to integer multiples of the grating's base frequency. The **`inc_field.f2`** and **`f3`** frequencies can be defined so that the multiple-order incident beams comprise a uniform frequency sampling over an extended range, as illustrated in Figure 40(b). Computational memory and runtime performance are greatly improved relative to including all incident frequencies in the zero order.

    Demo17 illustrates this approach for simulating diffraction of a Gaussian beam by a transmission grating. The beam transmits from glass to air, and the incidence medium is modeled as a semi-infinite medium of refractive index $n = 1.5$. Figure 41 illustrates the demo 17 graphic output showing the incident beam's far-field intensity profile, and Figure 42 shows the far-field diffracted intensity profile in all orders. The axes are scaled in direction-cosine units: $(\lambda/n) f_2$ and $(\lambda/n) f_3$ for the incident beam, λf_2 and λf_3 for the transmitted beam.
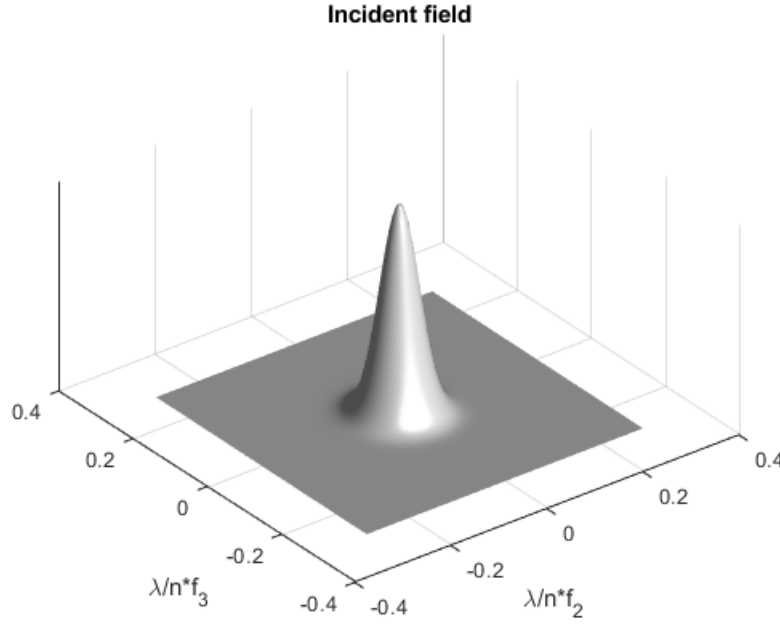
Figure 41. Incident beam's far-field intensity profile for demo 17.
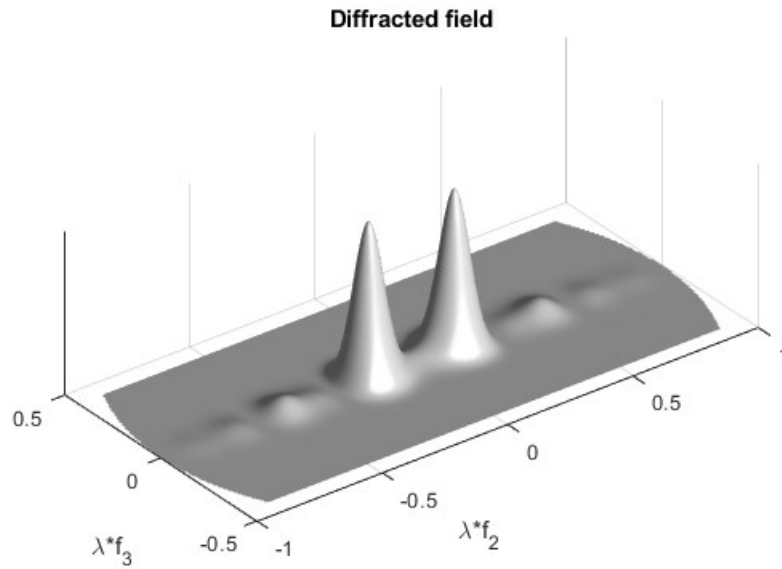
Figure 41. Diffracted beam's far-field intensity profile for demo 17.

The incident and transmitted beams' Fourier wave amplitudes are projected onto $\hat{s}$ and $\hat{p}$ polarization basis vectors, as described in GD-Calc.pdf, Eq's. 4.45 and 4.31,

$$ff\vec{E}_{m_1,m_2}^{[\text{Inc}]}[\vec{x}] = \hat{s}_{m_1,m_2}^{[\text{Inc}]}\, ffE_{\text{s},m_1,m_2}^{[\text{Inc}]}[\vec{x}] + \hat{p}_{m_1,m_2}^{[\text{Inc}]}\, ffE_{\text{p},m_1,m_2}^{[\text{Inc}]}[\vec{x}] \tag{32}$$

$$\vec{ff}E_{m_1,m_2}^{[\mathrm{T}]}[\vec{x}] = \hat{s}_{m_1,m_2}^{[\mathrm{T}]}\,ffE_{\mathrm{s},m_1,m_2}^{[\mathrm{T}]}[\vec{x}] + \hat{p}_{m_1,m_2}^{[\mathrm{T}]}\,ffE_{\mathrm{p},m_1,m_2}^{[\mathrm{T}]}[\vec{x}] \tag{33}$$

The linear mapping from incident to transmitted orders is described by GD-Calc.pdf, Eq. 4.47,

$$\begin{pmatrix} ffE_{\mathrm{s},m_1,m_2}^{[\mathrm{T}]}[\hat{e}_1\, b_1^{[0]}] \\ ffE_{\mathrm{p},m_1,m_2}^{[\mathrm{T}]}[\hat{e}_1\, b_1^{[0]}] \end{pmatrix} = \sum_{m_1',m_2'} \begin{pmatrix} T_{\mathrm{s,s},m_1,m_2,m_1',m_2'} & T_{\mathrm{s,p},m_1,m_2,m_1',m_2'} \\ T_{\mathrm{p,s},m_1,m_2,m_1',m_2'} & T_{\mathrm{p,p},m_1,m_2,m_1',m_2'} \end{pmatrix} \begin{pmatrix} E_{\mathrm{s},m_1',m_2'}^{[\mathrm{Inc}]}[\hat{e}_1\, b_1^{[L_1]}] \\ E_{\mathrm{p},m_1',m_2'}^{[\mathrm{Inc}]}[\hat{e}_1\, b_1^{[L_1]}] \end{pmatrix} \tag{34}$$

The correspondence between the quantities in the above equations and **gdc** data representations is outlined in GD-Calc.pdf, Table 4.1.

The Gaussian incident beam is assumed to have its waist at $x_1 = \hat{e}_1\, b_1^{[L_1]}$ (i.e., at the top grating surface). The beam's spatial amplitude profile at $x_1 = \hat{e}_1\, b_1^{[L_1]}$ is proportional to $\exp[-(x_2^2 + x_3^2)/w_0^2]$, where w_0 is the beam radius at the $1/e^2$ intensity level on the beam waist. The incident field expansion in GD-Calc.pdf, Eq. 4.39, is a Riemann-sum approximation to the beam's Fourier transform (plane-wave decomposition), which is proportional to $\exp[-\pi^2 (f_2^2 + f_3^2)\, w_0^2]$. The Fourier transform represents the far-field spatial amplitude mapping on a large sphere centered at $x_1 = \hat{e}_1\, b_1^{[L_1]}$ in the grating superstrate space. Direction cosines on the sphere (inward-directed) for incident order (m_1, m_2) are defined by the normalized wave vector $\vec{f}_{m_1,m_2}^{[\mathrm{Inc}]}\, \lambda / \sqrt{\varepsilon^{[\mathrm{sup}]}}$ (cf. GD-Calc.pdf, Eq. 4.42). The wave vectors are sampled on a uniform (f_2, f_3) grid (GD-Calc.pdf, Eq. 4.41), but are not uniformly sample on the sphere, so the summand in GD-Calc.pdf, Eq. 4.39 should include a denominator obliquity factor $OF_{m_1,m_2}^{[\mathrm{Inc}]}$ to account for the spherical integration weighting,

$$OF_{m_1,m_2}^{[\mathrm{Inc}]} = -f_{1,m_1,m_2}^{[\mathrm{Inc}]}\, \lambda / \sqrt{\varepsilon^{[\mathrm{sup}]}} \tag{35}$$

(The minus sign in Eq. 35 makes $OF_{m_1,m_2}^{[\mathrm{Inc}]}$ positive, GD-Calc.pdf, Eq. 4.42.)

Eq. 4.40 in GD-Calc.pdf is modified as follows to implicitly shift the x_1 origin to the top grating surface,

$$\vec{ff}E_{m_1,m_2}^{[\mathrm{Inc}]}[\vec{x}] = \vec{A}_{m_1,m_2}^{[\mathrm{Inc}]} \exp[i\, 2\pi\, \vec{f}_{m_1,m_2}^{[\mathrm{Inc}]} \bullet (\vec{x} - \hat{e}_1\, b_1^{[L_1]})] \tag{36}$$

With this definition, $\vec{A}_{m_1,m_2}^{[\mathrm{Inc}]}$ is the top-surface incident field amplitude: $\vec{A}_{m_1,m_2}^{[\mathrm{Inc}]} = \vec{ff}E_{m_1,m_2}^{[\mathrm{Inc}]}[\hat{e}_1\, b_1^{[L_1]}]$. The amplitudes $\vec{A}_{m_1,m_2}^{[\mathrm{Inc}]}$ are defined by three conditions: First, the incident beam is TE-polarized (i.e. $\vec{A}_{m_1,m_2}^{[\mathrm{Inc}]}$ is in the $\hat{e}_1$, $\hat{e}_3$ plane),

$$\hat{e}_2 \cdot \vec{A}^{[\text{Inc}]}_{m_1,m_2} = 0 \tag{37}$$

Second, the amplitude profile is a Gaussian divided by the obliquity factor,

$$\left| \vec{A}^{[\text{Inc}]}_{m_1,m_2} \right| = \exp[-\pi^2 \, ((f^{[\text{Inc}]}_{2,m_1,m_2})^2 + (f^{[\text{Inc}]}_{3,m_1,m_2})^2)\, w_0^{\,2}] \,/\, OF^{[\text{Inc}]}_{m_1,m_2} \tag{38}$$

Third, $\vec{A}^{[\text{Inc}]}_{m_1,m_2}$ is real-valued with a positive $\hat{e}_3$ projection, implying that all incident waves are in phase and the beam is in focus at $x_1 = \hat{e}_1 \, b_1^{[L_1]}$,

$$\text{Im}[\vec{A}^{[\text{Inc}]}_{m_1,m_2}] = \mathbf{0}, \quad \hat{e}_3 \cdot \vec{A}^{[\text{Inc}]}_{m_1,m_2} > 0 \tag{39}$$

$\vec{A}^{[\text{Inc}]}_{m_1,m_2}$ is projected onto $\hat{s}$ and $\hat{p}$ polarization bases (cf. GD-Calc.pdf, Eq. 4.45),

$$\vec{A}^{[\text{Inc}]}_{m_1,m_2} = \hat{s}^{[\text{Inc}]}_{m_1,m_2} \, A^{[\text{Inc}]}_{\text{s},m_1,m_2} + \hat{p}^{[\text{Inc}]}_{m_1,m_2} \, A^{[\text{Inc}]}_{\text{p},m_1,m_2} \tag{40}$$

Eq. 37 implies

$$\left(A^{[\text{Inc}]}_{\text{s},m_1,m_2} \quad A^{[\text{Inc}]}_{\text{p},m_1,m_2} \right) = u \left(p^{[\text{Inc}]}_{2,m_1,m_2} \quad -s^{[\text{Inc}]}_{2,m_1,m_2} \right) \tag{41}$$

for some scalar u. The left side of Eq. 38 equates to $\left| \vec{A}^{[\text{Inc}]}_{m_1,m_2} \right| = \sqrt{(A^{[\text{Inc}]}_{\text{s},m_1,m_2})^2 + (A^{[\text{Inc}]}_{\text{p},m_1,m_2})^2}$ (because basis vectors $\hat{s}^{[\text{Inc}]}_{m_1,m_2}$ and $\hat{p}^{[\text{Inc}]}_{m_1,m_2}$ in Eq. 40 are orthonormal), and substitution from Eq. 41 yields

$$u = \frac{\exp[-\pi^2 \, ((f^{[\text{Inc}]}_{2,m_1,m_2})^2 + (f^{[\text{Inc}]}_{3,m_1,m_2})^2)\, w_0^{\,2}]}{OF^{[\text{Inc}]}_{m_1,m_2} \, \sqrt{(p^{[\text{Inc}]}_{2,m_1,m_2})^2 + (s^{[\text{Inc}]}_{2,m_1,m_2})^2}} \tag{42}$$

The square root sign in Eq. 42 is chosen to satisfy Eq. 39,

$$\begin{aligned}
\hat{e}_3 \cdot \vec{A}^{[\text{Inc}]}_{m_1,m_2} &= u \left(s^{[\text{Inc}]}_{3,m_1,m_2} \, p^{[\text{Inc}]}_{2,m_1,m_2} - p^{[\text{Inc}]}_{3,m_1,m_2} \, s^{[\text{Inc}]}_{2,m_1,m_2} \right) \\
&= -u \, f^{[\text{Inc}]}_{1,m_1,m_2} \, \lambda \,/\, \sqrt{\varepsilon^{[\text{sup}]}} > 0
\end{aligned} \tag{43}$$

This result follows because the vector triplet $\left(\hat{s}^{[\text{Inc}]}_{m_1,m_2} \quad \hat{p}^{[\text{Inc}]}_{m_1,m_2} \quad \vec{f}^{[\text{Inc}]}_{m_1,m_2} \, \lambda \,/\, \sqrt{\varepsilon^{[\text{sup}]}} \right)$ forms a right-handed orthonormal basis set and $f^{[\text{Inc}]}_{1,m_1,m_2}$ is negative (GD-Calc.pdf, Eq's. 4.44 and 4.42).

The amplitudes $E^{[\mathrm{Inc}]}_{\mathrm{s},m_1',m_2'}[\hat{e}_1\,b_1^{[L_1]}]$ and $E^{[\mathrm{Inc}]}_{\mathrm{p},m_1',m_2'}[\hat{e}_1\,b_1^{[L_1]}]$ on the right side of Eq. 34 equate to $A^{[\mathrm{Inc}]}_{\mathrm{s},m_1,m_2}$ and $A^{[\mathrm{Inc}]}_{\mathrm{p},m_1,m_2}$ (based on Eq. 36). These amplitudes determine the transmitted field's Fourier amplitudes $A^{[\mathrm{T}]}_{\mathrm{s},m_1,m_2}$ and $A^{[\mathrm{T}]}_{\mathrm{p},m_1,m_2}$ from Eq 34,

$$A^{[\mathrm{T}]}_{\mathrm{s},m_1,m_2} = f\!f E^{[\mathrm{T}]}_{\mathrm{s},m_1,m_2}[\hat{e}_1\,b_1^{[0]}], \quad A^{[\mathrm{T}]}_{\mathrm{p},m_1,m_2} = f\!f E^{[\mathrm{T}]}_{\mathrm{p},m_1,m_2}[\hat{e}_1\,b_1^{[0]}] \tag{44}$$

$$\begin{pmatrix} A^{[\mathrm{T}]}_{\mathrm{s},m_1,m_2} \\ A^{[\mathrm{T}]}_{\mathrm{p},m_1,m_2} \end{pmatrix} = \sum_{m_1',m_2'} \begin{pmatrix} T_{\mathrm{s},\mathrm{s},m_1,m_2,m_1',m_2'} & T_{\mathrm{s},\mathrm{p},m_1,m_2,m_1',m_2'} \\ T_{\mathrm{p},\mathrm{s},m_1,m_2,m_1',m_2'} & T_{\mathrm{p},\mathrm{p},m_1,m_2,m_1',m_2'} \end{pmatrix} \begin{pmatrix} A^{[\mathrm{Inc}]}_{\mathrm{s},m_1,m_2} \\ A^{[\mathrm{Inc}]}_{\mathrm{p},m_1,m_2} \end{pmatrix} \tag{45}$$

The far-field transmitted amplitudes on a distant sphere include these amplitudes (applied to polarization basis vectors $\hat{s}^{[\mathrm{T}]}_{m_1,m_2}$ and $\hat{p}^{[\mathrm{T}]}_{m_1,m_2}$, Eq. 4.31 in GD-Calc.pdf), and also include an obliquity factor $OF^{[\mathrm{T}]}_{m_1,m_2}$ to account for the projected beam area in the propagation direction,

$$OF^{[\mathrm{T}]}_{m_1,m_2} = -\mathrm{Re}[f^{[\mathrm{T}]}_{1,m_1,m_2}]\lambda / \sqrt{\varepsilon^{[\mathrm{sub}]}} \tag{46}$$

This factor appears in the Kirchhoff integral for the far-field amplitude. The real part of $f^{[\mathrm{T}]}_{1,m_1,m_2}$ is used to mask evanescent orders (for which $f^{[\mathrm{T}]}_{1,m_1,m_2}$ is pure imaginary and $OF^{[\mathrm{IncT}]}_{m_1,m_2}$ is zero).

Appendix A. Algorithm notes for circle_partition.m

Figure A1 illustrates a unit circle whose first quadrant is partitioned into N blocks, e.g. $N = 4$ in Figure A1. The geometry is defined so that the shaded regions between the circle and the block profile all have the same area, denoted below as A. The block geometry is characterized in terms of the angles θ_1, θ_2, ... θ_{2N} at which the block profile intercepts the circle, and the angle sequence is constrained to have the following reversal symmetry,

$$\theta_j = \tfrac{\pi}{2} - \theta_{2N+1-j} \tag{A.1}$$
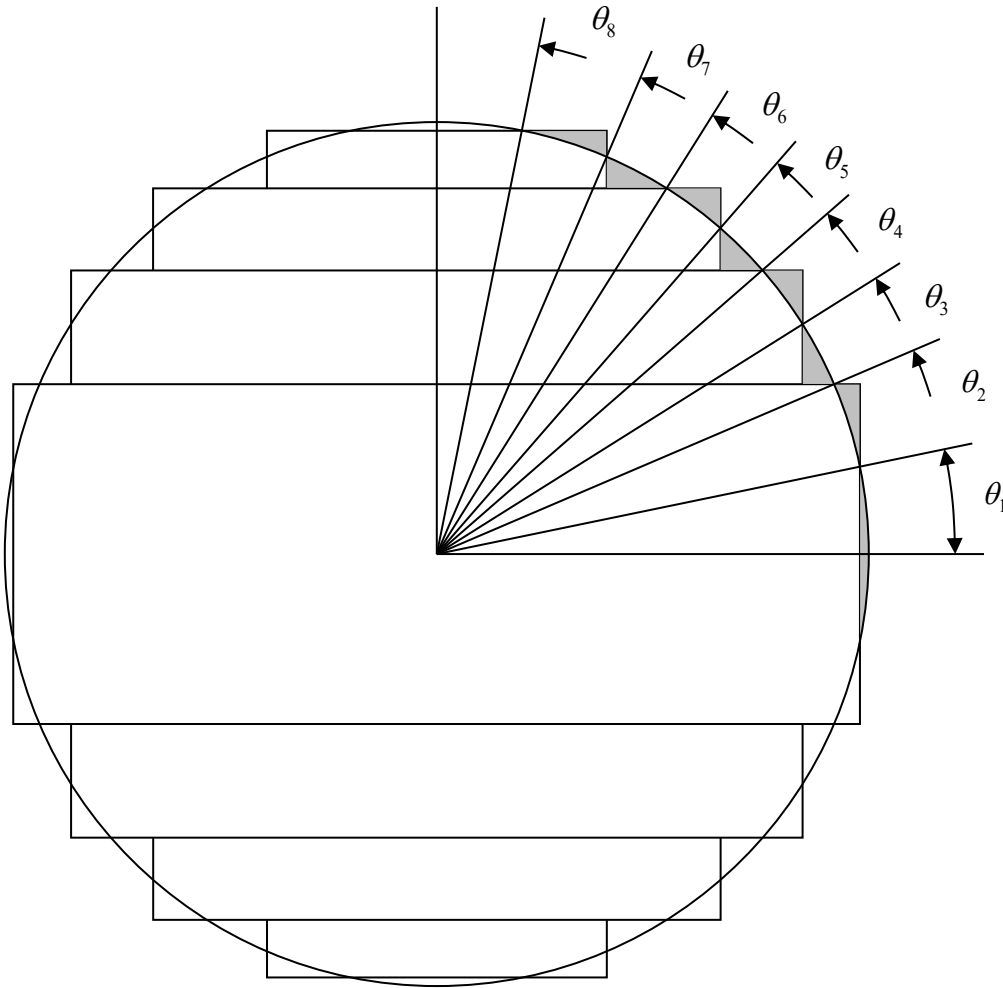

Figure A1. Block-partitioned circle geometry ($N = 4$).

The bottommost shaded region straddles two quadrants, and its area is

$$A = \theta_1 - \sin[\theta_1]\cos[\theta_1] \tag{A.2}$$

An analysis of the Figure A1 geometry yields the following relationship between successive intercept angles,

$$A = \begin{cases} \cos[\theta_j]\sin[\theta_{j+1}] - \frac{1}{2}(\cos[\theta_j]\sin[\theta_j] + \cos[\theta_{j+1}]\sin[\theta_{j+1}]) - \frac{1}{2}(\theta_{j+1} - \theta_j), & j \text{ odd} \\ \cos[\theta_{j+1}]\sin[\theta_j] - \frac{1}{2}(\cos[\theta_j]\sin[\theta_j] + \cos[\theta_{j+1}]\sin[\theta_{j+1}]) + \frac{1}{2}(\theta_{j+1} - \theta_j), & j \text{ even} \end{cases}$$

(A.3)

This implicitly represents a recursion relation defining θ_{j+1} from θ_j. (Newton's method is used to solve for θ_{j+1}.) The recursion is iterated for $j = 1\ldots N$, and the initial angle θ_1 is iteratively adjusted until the following terminal condition is satisfied:

$$\theta_{N+1} = \frac{\pi}{2} - \theta_N \qquad (A.4)$$

(from Eq. A.1).

The θ_1 iteration is initialized by making small-angle approximations in Eq. A.3. For $\theta_{j+1} - \theta_j \ll 1$, the following approximation is obtained,

$$A \cong \frac{1}{4}\sin[2\theta_j](\theta_{j+1} - \theta_j)^2 \qquad (A.5)$$

The angle difference $\theta_{j+1} - \theta_j$ can be approximated as a derivative,

$$\theta_{j+1} - \theta_j \cong \frac{d\theta}{dj} \qquad (A.6)$$

This results in the following approximation for j as a function of θ,

$$j \cong 1 + \int_{\theta_1}^{\theta_j} \sqrt{\frac{\sin[2\theta]}{4A}}\, d\theta \qquad (A.7)$$

The sine term is further approximated as

$$\sin[2\theta] \cong 2\theta \qquad (A.8)$$

so that Eq. A.7 reduces to

$$j \cong 1 + \frac{1}{3}\sqrt{\frac{2}{A}}\left(\theta_j^{3/2} - \theta_1^{3/2}\right) \qquad (A.9)$$

Hence,

$$\theta_j \cong \left((j-1)3\sqrt{\frac{A}{2}} + \theta_1^{3/2} \right)^{2/3} \tag{A.10}$$

The relationship between θ_1 and A defined by Eq. A.2 can be approximated as

$$A \cong \tfrac{2}{3}\theta_1^{\,3} \tag{A.11}$$

This substitution reduces Eq. A.10 to

$$\theta_j \cong \left((j-1)\sqrt{3} + 1 \right)^{2/3}\theta_1 \tag{A.12}$$

Eq. A.4 implies the following approximate equivalence,

$$\theta_{N+1/2} \cong \tfrac{\pi}{4} \tag{A.13}$$

Eq's. A.12 and A.13 yield the following initial approximation for θ_1,

$$\theta_1 \cong \tfrac{\pi}{4}\left((N-\tfrac{1}{2})\sqrt{3} + 1 \right)^{-2/3} \tag{A.14}$$

With this initialization, Newton's method (with a finite-difference derivative estimator) is used to eliminate the error in Eq. A.4 by successive refinement of the θ_1 estimate.

Appendix B. Achromatic prism design

Demo 12 is based on a grating design for achromatizing a transmission optical element such as a prism, Figure B1. A conventional, refracting prism is illustrated on the left (a), and an achromatic prism with a dispersion-compensating grating is shown on the right (b). The grating has a blazed, sawtooth profile with a facet angle of β. The prism's apex angle is α in configuration (a). In configuration (b) the apex angle relative to the grating substrate is $\alpha - \beta$, and relative to the facet surfaces is α.
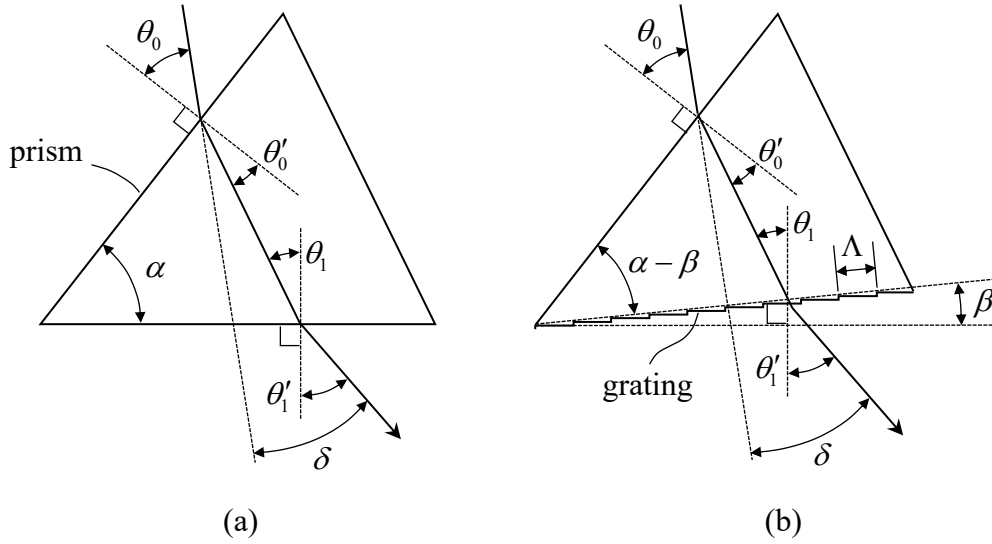
Figure B1. A conventional refracting prism (a) and a grating-compensated achromatic prism (b).

The ray angles illustrated in configuration (a) are determined by Snell's Law,

$$\sin\theta_0' = \frac{\sin\theta_0}{n} \tag{B.1}$$

$$\sin\theta_1' = n\sin\theta_1 \text{ (without the grating)} \tag{B.2}$$

where n is the refractive index and the θ's are ray angles defined relative to the surface normals. The ray deviation at the grating surface in configuration (b) is determined by the grating equation,

$$\sin[\theta_1' - \beta] = n\sin[\theta_1 - \beta] + \lambda/\Lambda \text{ (with the grating)} \tag{B.3}$$

where λ is the wavelength and Λ is the grating period. (The "$-\beta$" offset is applied in Eq. B.3 because θ_1 and θ_1' are defined relative to the facet surface normal, whereas the grating equation applies to the angles relative to the substrate normal.) However, Snell's Law, Eq. B.2, also applies on the grating surface at the blaze wavelength. (The equivalence between the grating equation and Snell's Law essentially defines the blaze wavelength, at which diffraction efficiency is maximized.)

Based on the geometry of configuration (a), the apex angle α and the total ray deviation angle δ have the following relationships to the ray angles,

$$\alpha = \theta_0' + \theta_1 \tag{B.4}$$

$$\delta = \theta_0 - \theta_0' + \theta_1' - \theta_1 \qquad (B.5)$$

Demo 12 assumes a symmetric ray geometry in configuration (a) at a particular design wavelength,

$$\theta_1 = \theta_0' \qquad (B.6)$$

Eq's. B.1, B.2, B.4, B.5, and B.6 are combined to eliminate the four ray angles and solve for α as a function of δ and n,

$$\alpha = 2\,\text{atan}\left[\frac{\sin[\delta/2]}{n - \cos[\delta/2]}\right] \qquad (B.7)$$

n is implicitly a function of λ and Eq. B.7 is applied at the design wavelength.

For configuration (b), the same apex angle α and incidence angle θ_0 are used to keep the ray geometry unchanged at the blaze wavelength (which is the ray trace design wavelength). β and Λ are determined by two conditions: First the grating equation is equivalent to Snell's Law (i.e., Eq. B.3 holds with θ_1' defined by Eq. B.2). Second, the derivative of δ (as determined by Eq. B.3) with respect to λ should be zero to achieve narrow-band achromaticity. The following derivative relations are obtained from Eq's. B.1, B.4, B.3, and B.5, and the condition $d\delta/d\lambda = 0$:

$$\frac{d\theta_0'}{d\lambda}\cos\theta_0' = -\frac{dn}{d\lambda}\frac{\sin\theta_0}{n^2} \qquad (B.8)$$

$$\frac{d\theta_1}{d\lambda} = -\frac{d\theta_0'}{d\lambda} \qquad (B.9)$$

$$\frac{d\theta_1'}{d\lambda}\cos[\theta_1' - \beta] = \frac{dn}{d\lambda}\sin[\theta_1 - \beta] + n\frac{d\theta_1}{d\lambda}\cos[\theta_1 - \beta] + 1/\Lambda \qquad (B.10)$$

$$\frac{d\delta}{d\lambda} = -\frac{d\theta_0'}{d\lambda} + \frac{d\theta_1'}{d\lambda} - \frac{d\theta_1}{d\lambda} = 0 \qquad (B.11)$$

Eq's. B.4, B.5, B.9 and B.11 are used to eliminate θ_1, θ_1', and their derivatives in Eq's. B.3 and B.10,

$$\sin[\delta + \alpha - \theta_0 - \beta] = n\sin[\alpha - \theta_0' - \beta] + \lambda/\Lambda \qquad (B.12)$$

$$0 = \frac{dn}{d\lambda}\sin[\alpha - \theta_0' - \beta] - n\frac{d\theta_0'}{d\lambda}\cos[\alpha - \theta_0' - \beta] + 1/\Lambda \qquad (B.13)$$

Λ is eliminated between Eq's. B.12 and B.13, and the result is solved for β,

$$\beta = \operatorname{atan}\left[\frac{\left(\lambda\dfrac{dn}{d\lambda}-n\right)\sin\left[\alpha-\theta_0'\right]-n\,\lambda\dfrac{d\theta_0'}{d\lambda}\cos\left[\alpha-\theta_0'\right]+\sin\left[\delta+\alpha-\theta_0\right]}{\left(\lambda\dfrac{dn}{d\lambda}-n\right)\cos\left[\alpha-\theta_0'\right]+n\,\lambda\dfrac{d\theta_0'}{d\lambda}\sin\left[\alpha-\theta_0'\right]+\cos\left[\delta+\alpha-\theta_0\right]}\right]$$

(B.14)

Figure B2 illustrates the prism's dispersion characteristic (δ vs λ). The Demo 12 design is based on an acrylic (PMMA) prism material. This type of grating structure could be used, for example, to achromatize a Fresnel lens.
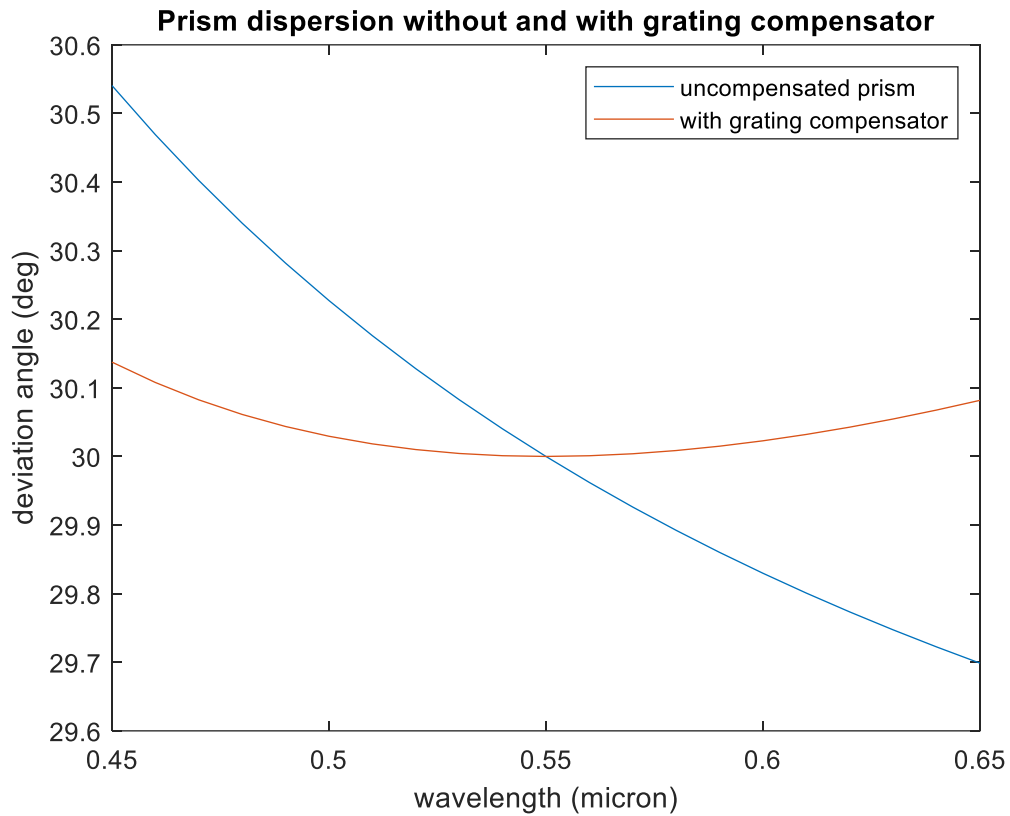

Figure B2. Deviation angle δ versus wavelength λ for a prism without and with achromatization.