The modal method : a reference method for modeling of the 2D metal diffraction gratings

Ivan Gushchin

Université de Saint-Etienne - Jean MONNET
Ecole doctorale Science Ingénierie Santé

THÈSE

Pour obtenir le grade de docteur en sciences
Spécialité : Optique, optoelectronique, photonique.

La méthode modale: une méthode de référence pour la modélisation de réseaux de diffraction métalliques deux dimensionnel

par:
Ivan GUSHCHIN

a été soutenue le 12 juillet 2011, devant le jury composé de:

Patrice BALDECK	Université Joseph Fourier, Grenoble	Examinaeur
Gérard GRANET	Université Blaise Pascal, Clermont-Ferrand 2	Rapporteur
Brahim GUIZAL	Université Montpellier 2	Examinateur
Olivier PARRIAUX	Université Jean Monnet, Saint-Etienne	Président
Alexandre Tishchenko	Université Jean Monnet, Saint-Etienne	Directeur de thèse

University of Saint-Etienne - Jean MONNET
Graduate School of Health, Science, Engineering

THESIS

to get the Doctor of Science degree
Speciality: Optics, optoelectronic, photonic.

The modal method: a reference method for modeling of the 2D metal diffraction gratings

by:
Ivan GUSHCHIN

was supporteded on the 12 of July 2011, before the committee:

Patrice BALDECK	University Joseph Fourier, Grenoble	Examinator
Gérard GRANET	University Blaise Pascal, Clermont-Ferrand 2	Rapporteur
Brahim GUIZAL	University Montpellier 2	Examinator
Olivier PARRIAUX	University Jean Monnet, Saint-Etienne	Presindent
Alexandre Tishchenko	University Jean Monnet, Saint-Etienne	Supervisor of the thesis

Résumé

Les éléments de diffraction sont largement utilisés aujourd'hui dans un nombre grandissant d'applications grâce à la progression des technologies de microstructuration dans le sillage de la microélectronique. Pour un design optimal de ces éléments, des méthodes de modélisation précises sont nécessaires. Plusieurs méthodes ont été développées et sont utilisées avec succès pour des réseaux de diffraction unidimensionnel de différents types. Cependant, les méthodes existantes pour les réseaux deux dimensionnel ne couvrent pas tous types de structures possibles. En particulier, le calcul de l'efficacité de diffraction sur les réseaux métalliques à deux dimensionnel avec parois verticales représente encore une grosse difficulté pour les méthodes existantes. Le présent travail a l'objectif le développement d'une méthode exacte de calcul de l'efficacité de diffraction de tels réseaux qui puisse servir de référence.

La méthode modale développée ici - dénommée ,,true-mode" en anglais - exprime le champ électromagnétique sur la base des vrais modes électromagnétiques satisfaisant les conditions limites de la structure 2D à la différence d'une méthode modale où les modes sont ceux d'une structure approchée obtenue, par exemple, par développement de Fourier. L'identification et la représentation de ces vrais modes à deux dimensions restait à faire et ce n'est pas le moindre des résultats du présent travail que d'y avoir conduit.

Les expressions pour la construction du champ sont données avec des exemples de résultats concrets. Sont aussi fournies les équations pour le calcul des intégrales de recouvrement et des éléments de la matrice de diffusion.

Resume

Diffractive elements are widely used in many applications now as the microstructuring technologies are making fast progresses in the wake of microelectronics. For the optimization of these elements accurate modeling methods are needed. There exists well-developed and widely used methods for one-dimensional diffraction gratings of different types. However, the methods available for solving two-dimensional periodic structures do not cover all possible grating types. The development of a method to calculate the diffraction efficiency of two-dimensional metallic gratings represents the objective of this work.

The one-dimensional true-mode method is based on the representation of the field inside the periodic element as a superposition of particular solutions, each one of them satisfying exactly the boundary conditions. In the developed method for the two-dimensional gratings the representation of the field within the grating in such way is used.

In the present work, the existing modal methods for one-dimensional gratings can be used as the basis for the construction of the modal field distribution functions within two-dimensional gratings. The modal function distributions allow to calculate the overlap integrals of the fields outside the grating with those within the structure. The transition matrix coefficients are formed on the basis of these integrals. The final stage is the calculation of the scattering matrix based on two transition matrices.

The equations for the field reconstruction are provided and accompanied by examples of results. Further equations used to calculate the overlap integrals and scattering matrix coefficients are provided.

Résumé substantiel

Les éléments de diffraction sont largement utilisés aujourd'hui dans un nombre grandissant d'applications grâce à la progression des technologies de microstructuration dans le sillage de la microéilectronique. Pour un design optimal de ces éléments, des méthodes de modélisation précises sont nécessaires. Plusieurs méthodes ont été développées et sont utilisées avec succès pour des réseaux de diffraction unidimensionnel de différents types. Cependant, les méthodes existantes pour les réseaux deux dimensionnel ne couvrent pas tous types de structures possibles. En particulier, le calcul de l'efficacité de diffraction sur les réseaux métalliques à deux dimensionnel avec parois verticales représente encore une grosse difficulté pour les méthodes existantes (méthodes FDTD, Fourier-modal). Le présent travail a l'objectif le développement d'une méthode exacte de calcul de l'efficacité de diffraction de tels réseaux qui puisse servir de référence.

La méthode modale développée ici - dénommée ,,true-mode" en anglais - exprime le champ électromagnétique sur la base des vrais modes électromagnétiques satisfaisant les conditions limites de la structure 2D à la différence d'une méthode modale où les modes sont ceux d'une structure approchée obtenue, par exemple, par développement de Fourier. L'identification et la représentation de ces vrais modes à deux dimensions restait à faire et ce n'est pas le moindre des résultats du présent travail que d'y avoir conduit.

Le rappel des notions de base est fait dans l'introduction à partir des équations de Maxwell dont on va chercher les solutions satisfaisant les conditions limites pour résoudre le problème de la diffraction. Le formalisme utilisé pour l'expression d'ondes planes incidentes de polarisation quelconque dans un espace homogène est établi ainsi que la forme utilisée pour l'expression des résultats intermédiaires et finaux. Cette introduction est supposée aider ceux qui sont intéressés à travailler sur la méthode modale.

Dans ce travail, les modes des réseaux à une dimension sont utilisés comme base pour la construction du champ des modes des réseaux à deux dimensions. Pour cette raison, l'introduction au chapitre suivant est consacrée à la méthode modale de Fourier ainsi qu'à la méthode des vrais modes qui

y sont décrites en détail. Dans la section sur la méthode des vrais modes, le lecteur peut identifier les étapes principales conduisant à la résolution du problème de diffraction.. Cette partie est un extrait de l'état des connaissances existantes dans le cas unidimensionnele. Le lecteur est renvoyé aux sources primaires mais il pourrait déjà développer son propre code à une dimension sur la base du contenu de cette section.

La méthode modale de Fourier est ensuite décrite avec une extension originale pour son application à des profils à parois inclinées. Elle peut être utilisée avec profit pour la représentation du champ dans chaque tranche élémentaire en lesquelles le profil réel est segmenté lorsque les bords sont notablement obliques. Lorsqu'ils sont verticaux le choix de la méthode des vrais modes est plus indiqué.

La partie centrale de ce travail est divisée en deux sections. La première est consacrée aux modes à deux dimensions latérales: définition du mode, représentation de sa forme, recherche du mode et de son champ; la deuxième section contient la description des étapes ultérieures nécessaires pour aboutir à la matrice de diffusion. Toutes les étapes coïncident avec celles de la méthode consacrée aux réseaux à une dimension. Après que les constantes de propagation ont été obtenues le champ modal est construit et un coefficient de normalisation est déterminé. La connaissance du champ modal permet de calculer les intégrales de recouvrement aux frontières du réseau à deux dimensions. La dernière étape est le calcul de la matrice de diffusion à partir des matrices de transition de chaque tranche du profil.

Les expressions pour la construction du champ sont données avec des exemples de résultats concrets. Sont aussi fournies les équations pour le calcul des intégrales de recouvrement et des éléments de la matrice de diffusion.

Les mots clés

- diffraction
- réseaux de diffraction métalliques
- méthode modale
- réseaux deux dimensionnel

The key words

- diffraction
- metal diffraction gratings
- modal method
- two-dimensional gratings
- true-modal

Cette thèse de doctorat a été préparée au:

Laboratoire Hubert CURIEN
UMR CNRS 5516
18 rue du Pr. Benoît LAURAS
42000 Saint-Etienne
France

http://laboratoirehubertcurien.fr/

Contents

Chapter 1

Preface

1.1 Short problematics

The solution of the diffraction problem answers, in general, to the question what light distribution or what field distribution will there be after the interaction with the diffraction element. When the problem is formulated in this way there is no difference or no interest in the processes which occur inside the grating. Chandezon's method[1], Rayleigh method[2], Fourier-modal method[3] are the methods of this kind: they answer the question what is the light in the reflection region and what light is transmitted.

This answer is enough for the most variety of problems. However, these methods can not explain why there certain effect occurs. For example, Chang-Hasnain[4] and Bonnet[5] effects can not be explained with the methods mentioned above. To explain the phenomena and to interpret the behaviour of the grating it is necessary to "look inside" the grating and to understand what happens inside the grating.

Diffraction gratings found theirs application in the laser technology. They are used for selecting the polarisation inside the cavity, for the spectral selectivity of the induced light, etc. The use of gratings in the lasers with increasing output power meet with the need to calculate the ultimate field density inside the laser mirrors. After a certain threshold exceeding, a thermally induced geometry change of the mirror begins. If the power dissipation

is not enough, physical destruction of the mirror begins and the laser mirror usage becomes impossible. Mirrors can be also damaged with effects induced with the high density of the electromagnetic field. The modal method is the most suitable one to answer the questions: what is the field inside the grating, what is the electromagnetic field distribution. The modal method can respond to the questions posted above. The solution of the diffraction problem is built in the modal method in the following steps: a set of solutions to Maxwell equations is constructed so that each solution (mode) satisfies the boundary conditions imposed by the grating and the periodicity conditions imposed by the incident light; parts of the incident radiation transmitted into each mode are calculated on the next step; the modes propagate to the opposite grating board with with theirs propagation constants carrying theirs energy portion; at the opposite border, energy of the modes is partially reflected backwards and is partially transmitted into the semi-infinite space at the opposite side of the grating. It is worth noting that the reflected modes, after they reach the initial boundary, are also partially reflected and partially transmitted into the media of incidence. The process is more complicate but this description is appropriate for the review of the processes occurring inside the grating.

1.2 The interest in the two-dimensional gratings

To meet the requirements and to be in time with the society's demands it is necessary not only to improve the technological methods but it is also necessary to have new methods of calculation and assessment of manufactured materials. For definite applications,the methods already developed fail to satisfy new requirements. For example, the decrease of operational wavelength leads to failure of the existing method in one-dimensional grating modelling. New materials require more accurate methods. Currently, two-dimensional grating are not used as widely as one-dimensional gratings. One-dimensional gratings have found their place in wide spectra of applications: reflectors[6, 4]

and ultra-broadband mirrors[7], encryption[8], beam-splitters[9], diodes[10], coronagraphy[11], sensors[12]. Two-dimensional gratings are also applied in many fields, like optical demultiplexers[13], solar cells[14], photonic crystals theory[15], plasmonics[16].

The most convincing application of the two-dimensional gratings arises from the diffracted light independence (or smooth dependence) on the incident light polarisation. Another application can be found in checking of the measured results along one direction with respect to the results obtained along another direction. Using of the two-dimensional grating for such application allows to remain the sample untouched. Currently, there are methods to calculate two-dimensional grating. They have drawbacks but they are already used to obtain or estimate diffraction efficiencies.

The development of another method is needed for comparison of the results obtained by one method with the results obtained by another method. A high accuracy of the results together with calculation speed are the requirements for a new method. A capability to deal with any kind of materials is also one of the requirements.

1.3 Mode. State

The nature of the light and related effects excite people for a long time. A stained glass, mirrors, Christmas ornaments proves this interest. A craving for diamonds, which has delicious play of the light inside, is another example of a large human interest in things, which transform and convert the light.

The diffraction theory of light on the periodic structures does not have a history as long as craving for diamonds, but periodic structures as optical elements have found their place really quickly. These elements are used today in many technological and production areas. Meanwhile, the use of the properties of periodic structures for each specific application requires more accurate results and deep understanding of the light interaction with the structure. Today a high accuracy of the diffraction elements characteristics is well demanded and new method are required.

Despite the fact that the problem is simple at first glance, up to present

days a development of methods for solving the diffraction problem continues. Each of them finds its adherents, developers and opponents. The Differences between the methods are in the basic principles, which are addressed while a method development that leads to the number of methods, so each of them is more optimal for a particular grating profile.

Undoubtedly that each method solves the same equations with the same boundary conditions. But when one comes close to the particular problem he has to make some assumptions or reduce the model to some cases.

Quantum physics introduced a concept of the state. The state is one of possible configurations, one of possible solutions. It is something that can characterise the system. Each solution of the quantum physics problem is a superposition of the problem's states.

A mode concept is similar to the state one. The mode is some characteristic of the system. Some kind of the general problem solution which presents in any particular problem solution. We will refer to that abstract definition of the mode further in this work. The mode is something that posses certain properties. The mode is characterised by it's properties but not by some physical action or a phenomenon.

1.4 Structure of the thesis

The following sequence of the presentation is found optimal for this work. First part is devoted to the introduction of the mathematical background referred in the rest of the work. Basic concepts are expressed so that each person willing to get into this work is able to find all necessary information in one place. The literature review provides brief description of the methods which are commonly and widely used for the diffraction problem resolution. Their advantages and weak points in comparison with the modal methods are marked.

Modal methods are described separately and in details because they are placed as a basis for the modal method on the two-dimensional diffraction gratings. In these chapter reader will find generalisation of the research in the modal methods up to date. Provided information is sufficient for the

separated introduction to the one-dimensional modal methods.

The next chapter is the central part of the work. The modal method to calculate diffraction efficiency of the two-dimensional grating is developed. All the traditional steps for the modal methods are described. The chapter begins with the problem formulation, continue to the dispersion equation and the dispersion function construction. The following step shows how to get propagation constants and modal fields' distributions. The overlap integrals are then calculated on the basis of the modal field distribution. They are used in the transition matrix on the following step. And the scattering matrix composition is the final resolution step.

The conclusion chapter summaries all the information from the previous chapters.

1.5 Activity report for all the thesis period

I also want to point out what achievements were reached during the thesis work. First step was to investigate state of the art for the moment of the thesis beginning. Main efforts were devoted to the modal vision, understanding of the modal concept. The literature can provide description of the concept, but the feeling, the inside of the method and of the modal approach can not be expressed in the papers. I have followed the founders of the method from the first publication to the present. Each publication gave new area where method can be applied or where it became possible to use the method.

The familiarisation with the Rigorous Coupled Wave Analysis (RCWA) was necessary as it is fraternal method for the true-modal method. RCWA method is also called Fourier-modal method, because the modal decomposition in Fourier space is the key property of the method. Nevière and Popov have improved their differential method by taking into account the inclined boundaries in [17]. This idea was successfully realised for the Fourier modal method and it was shown that application of this technique leads to the substantial accuracy improvement with respect to the traditional implementation of the RCWA.

An attempt to apply this technique to the true-modal method has not suc-

ceeded. I believe that this idea can not be applied to the true-modal method due to the properties of the method. The equation system for the true-modal method is defined completely. I have not found possibility to introduce new unknown or new characteristic with unknown amplitude. All the attempts lead either to the over-defined equation systems (and the introduced parameter did not influence on the results, unknown could be neglected) or to the solution out of the definition zone (fields were with constant term, so that solution was not applicable for the oscillating field (electromagnetic field).

The tilting of the all coordinate system for the true-modal method has been already published [18]. This approach has shown its efficiency for the lamellar tilted gratings. A success of the idea application, with regard to the fail in the slanted boundaries account, can be attributed to the fact that in the tilted coordinate system the problem became traditional for the modal method: with the walls, orthogonal to the internal waves.

After I became on a short hand with the modal method, I moved towards developing of the true-modal method for the two-dimensional structures. The problem was initially reduced to the structures with orthogonal periodicity vectors. The theoretical part of this method is similar to the one-dimensional ancestor. The two-dimensional case introduces specific properties like coupled modes even for dielectric gratings. The main difficulty was to implement and debug the realisation of the method. Problem arises from the limited number of reference methods which can be used on each step of the development process. RCWA method provides propagation constants and diffraction efficiencies in the far-field of the dielectric structure. All the step between these two points are well hidden and can not be used for analysis purposes.

The two-dimensional method is developed. The results obtained with this method can be used for the approximate methods calibration.

Chapter 2

Introduction

2.1 Maxwell's equations

Maxwell's equations are a set of partial differential equations that, together with the Lorentz force law, form the foundation of classical electrodynamics, classical optics, and electronic circuits. These in turn underlie modern electrical and communications technologies. [1]

Maxwell's equations describe conceptually how electric charges and electric currents act as sources of the electric and magnetic fields. Furthermore, they describe how a time varying electric field generates a time varying magnetic field and vice verse. Among the four equations, two Gauss's law describe how the fields emanate from charges. (For the magnetic field there is no magnetic charge and therefore magnetic fields lines neither begin nor end anywhere.) The other two equations describe how the fields 'circulate' around their respective sources; the magnetic field 'circulates' around electric currents and time varying electric field in Ampère's law with Maxwell's correction, while the electric field 'circulates' around time varying magnetic fields in Faraday's law.

Gauss's law

$$\vec{\nabla} \cdot \vec{\mathbf{D}} = \rho$$

[1] This section is partial copy of the article, which can be found on the [19] and further links.

where $\vec{\nabla}\cdot$ denotes divergence, $\vec{\mathbf{D}}$ is the electric displacement field, and ρ is the total electric charge density (including both free and bound charge).

Gauss's law describes the relationship between an electric field and the generating electric charges: the electric field points away from positive charges and towards negative charges. In the field line description, electric field lines begin only at positive electric charges and end only at negative electric charges. 'Counting' the number of field lines in a closed surface, therefore, yields the total charge enclosed by that surface. More technically, it relates the electric flux through any hypothetical closed "Gaussian surface" to the electric charge within the surface.

Gauss's law for magnetism

$$\vec{\nabla} \cdot \vec{\mathbf{B}} = 0$$

where $\vec{\nabla}\cdot$ denotes divergence and $\vec{\mathbf{B}}$ is the magnetic induction.

Gauss's law for magnetism states that there are no "magnetic charges" (also called magnetic monopoles), analogous to electric charges. Instead, the magnetic field due to materials is generated by a configuration called a dipole. Magnetic dipoles are best represented as current loops but resemble positive and negative 'magnetic charges', inseparably bound together, having no net 'magnetic charge'. In terms of field lines, this equation states that magnetic field lines neither begin nor end but make loops or extend to infinity and back. In other words, any magnetic field line that enters a given volume must somewhere exit that volume. Equivalent technical statements are that the sum total magnetic flux through any Gaussian surface is zero, or that the magnetic field is a solenoidal vector field.

Faraday's law

$$\vec{\nabla} \times \vec{\mathbf{E}} = -\frac{\partial \vec{\mathbf{B}}}{\partial t}$$

where $\vec{\nabla}\times$ denotes curl and $\vec{\mathbf{B}}$ is the magnetic induction.

Faraday's law describes how a time varying magnetic field creates ("induces") an electric field. This aspect of electromagnetic induction is the operating principle behind many electric generators: for example a rotating

bar magnet creates a changing magnetic field, which in turn generates an electric field in a nearby wire.

Ampère's law with Maxwell's correction

$$\vec{\nabla} \times \vec{\mathbf{H}} = \vec{\mathbf{j}} + \frac{\partial \vec{\mathbf{D}}}{\partial t}$$

where $\vec{\nabla}\times$ denotes curl, $\vec{\mathbf{H}}$ is the magnetising field, $\mathbf{j}$ is the current density, and $\vec{\mathbf{D}}$ is the electric displacement field.

Ampère's law with Maxwell's correction states that magnetic fields can be generated in two ways: by electrical current (this was the original "Ampère's law") and by changing electric fields (this was "Maxwell's correction").

Maxwell's correction to Ampère's law is particularly important: It means that a changing magnetic field creates an electric field, and a changing electric field creates a magnetic field. Therefore, these equations allow self-sustaining "electromagnetic waves" to travel through empty space (see electromagnetic wave equation).

The speed calculated for electromagnetic waves, which could be predicted from experiments on charges and currents, exactly matches the speed of light; indeed, light is a form of electromagnetic radiation (so are X-rays, radio waves, and others). Maxwell understood the connection between electromagnetic waves and light in 1861, thereby unifying the previously-separate fields of electromagnetism and optics.

The light is an electro-magnetic wave with wavelength belonging to a definite range (human visible light for example has wavelengths ranging from $760nm$ to $380nm$). Maxwell's equations are valid for all the electro-magnetic waves including light.

In case of the light propagation we can exclude displacement current $\mathbf{j}$ and electrical charges density ρ from the equations. The system is resolved (as a rule) for materials with permeability equals that of vacuum ($\mu = \mu_0$). The wavelength is fixed for particular problem formulation and light supposed to be monochromatic. The time dependence thus far is described with $\exp(-j\omega t)$ term. We can rewrite the original system, taking in account

previous assumptions, in the form:

$$\begin{aligned} \vec{\nabla} \times \vec{\mathbf{E}} &= j\omega\mu_0\vec{\mathbf{H}} \\ \vec{\nabla} \times \vec{\mathbf{H}} &= -j\omega\varepsilon\vec{\mathbf{E}} \\ \vec{\nabla} \cdot \vec{\mathbf{D}} &= 0 \\ \vec{\nabla} \cdot \vec{\mathbf{H}} &= 0 \end{aligned} \tag{2.1}$$

The resolution of the system (2.1) with given boundary conditions gives the diffraction problem solution.

2.1.1 Incident light polarisation

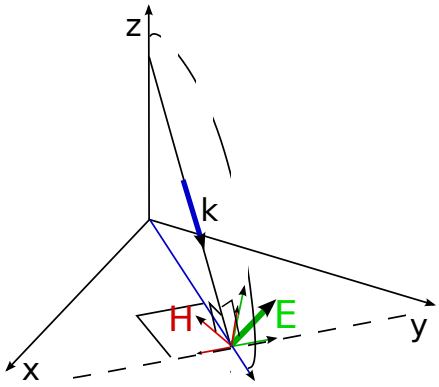


Figure 2.1: Illustration of the electromagnetic wave decomposition

Vectors of electric and magnetic fields of a plane wave in an isotropic medium are mutually orthogonal. Their amplitudes are strictly related. Any of these vectors in turn can be represented as a superposition of two orthogonal vectors. This allows to represent a plane wave as a superposition of two orthogonally polarised waves propagating in the same direction.

Plane of incidence of a plane wave on some boundary is defined so that both vectors, one is normal to the boundary and another is the wave vector of incidence wave, belong to that plane of incidence.

A wave with magnetic field lying in the plane of incidence and electric field orthogonal to this plane is called the TE-wave. Wave, which magnetic field is orthogonal to the plane of incidence is called TM-wave.

Decomposition of the incident electromagnetic wave with an arbitrarily oriented electric field (and associated magnetic field) into two differently polarised waves is unique, simplifies expression of fields and allows to manage all polarisation states.

2.1.2 Vector diagram of the diffracted light

Knowing the grating periods allows us to make assumptions about what orders will be propagated (to be measurable in the far field) but does not allow to define the intensity of each of the orders. As an example of such expectations on the period would have resulted in a grating with a period smaller than the wavelength in the case of normal incidence: reflection occurs only in the zeroth order, while regarding to the substrate material there no transmission can occur into the zeroth as well as into higher diffraction orders. The expression for the angles at which diffraction maxima will be visible is the following:

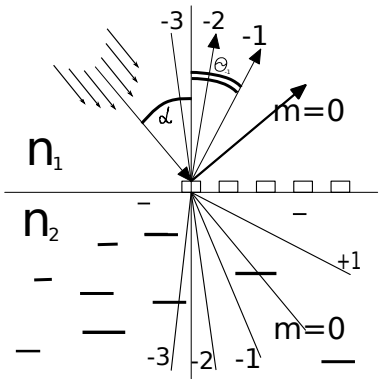


Figure 2.2: Diffraction notation

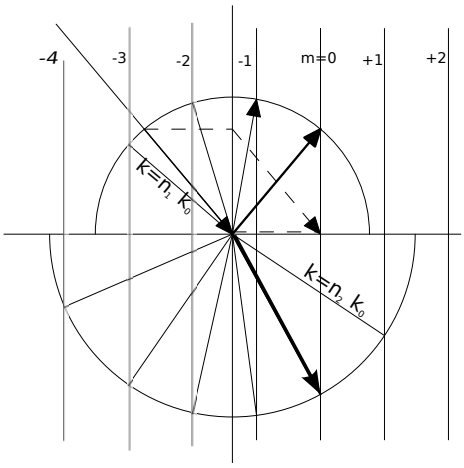
$$\Lambda \left(n_1 \sin \alpha - n_2 \sin \theta \right) = m\lambda$$

where m denotes the diffraction order number, n_1 - refractive index of the media, from which the light insides, n_2 is the refractive index of the media into which the waves are transmitted (is case of reflected waves $n_2 = n_1$), α is the incidence angle, Θ is the diffraction maximum angle , Λ is the grating period. From this expression, given that $\sin x \leq 1$ easily get the condition under which the diffraction peaks are visible, i.e. order will be observed in the far field:

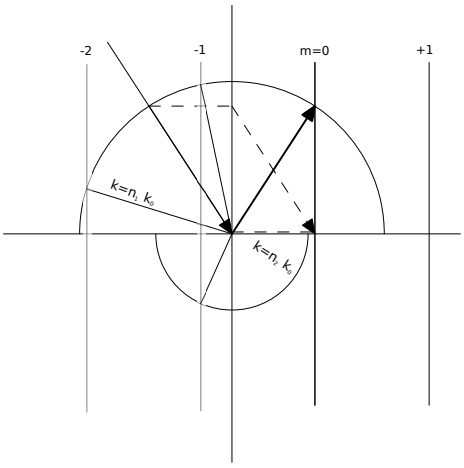
$$\left| \frac{n_1 \sin \alpha - \frac{m\lambda}{\Lambda}}{n_2} \right| \leq 1 \tag{2.2}$$

Intuitively, the technique of vector diagrams demonstrates this passage. The essence of this technique is that represented by two semicircles with a common centre on both sides of the horizontal line. On one side is represented by semicircle with a radius equal to $k_1 = \frac{2\pi}{\lambda_0} n_1$, i.e., value of the wave vector in the half with coefficient index n_1, and on the other hand, respectively, $k_2 = \frac{2\pi}{\lambda_0} n_2$, that is equal to the wave vector in the second half. Then, depicted

a line from the centre These semi-circles in the direction of the corresponding angle of incidence. Parallel segment connecting the centre and a point on the semicircle is transferred so that the point of the semicircle hit on the normal to the horizontal line drawn through the centre of a semicircle. The point where to move initially at the centre of semi circle, will reference and match zero directions. Through this point of the line perpendicular to the horizontal, we get 2 points of intersection with the semicircle. Having line from the centre of semi-circles to these points of intersection, we clearly see at what angle will apply the zeroth order. In order to see other orders, we put both sides of the reference points of segments equal to $K = \frac{2\pi}{\Lambda}$. Restoring



a) Example of the configuration when substrate is with larger refractive index then the cover of the grating. Grating vector is small with respect to both the indexes.



b) Example of the configuration when substrate is with small refractive index. Grating vector is smaller then the wave vector in the cover but larger then in substrate.

from each such point the normal to the horizontal line, we get the point of intersection a semicircle. Carrying out these points from the centre line of half circles, we get directions distribution of diffraction orders. The number of diffraction order corresponds to the number of the line and the sign of the order's number corresponds to the the direction in which we wove aside the zeroth reflected order (if in the same direction, where the source was moved segment corresponding to the incident wave - positive, the opposite - negative). In the case when the intersection with the semicircle does not occur,

the diffracted waves do not propagate, and condition (2.2) fails.

2.1.3 Littrow configuration. Brewster's angle

A special but common case is that in which the light is diffracted back towards the direction from which it came ($\alpha = -\beta$); this is called *Littrow configuration*, for which the grating equation becomes

$$m\lambda = 2d\sin(\alpha)$$

Brewster's angle (also known as the polarisation angle) is an angle of incidence at which the TM-polarised light is perfectly transmitted through a transparent dielectric surface, with no reflection. When unpolarised light is incident at this angle, the light that is reflected from the surface is therefore perfectly linearly TE-polarised. This special angle of incidence is named after the Scottish physicist, Sir David Brewster (1781–1868).

2.1.4 Poynting vector

The Poynting vector can be thought of as representing the energy flux of an electromagnetic field. It is named after its inventor John Henry Poynting. Oliver Heaviside and Nikolay Umov independently co-invented the Poynting vector. In Poynting's original paper and in many textbooks it is defined as

$$\vec{\mathbf{S}} = \vec{\mathbf{E}} \times \vec{\mathbf{H}}$$

which is often called the Abraham form; hereis $\vec{\mathbf{E}}$ the electric field and $\vec{\mathbf{H}}$ the auxiliary magnetic field. Sometimes, an alternative definition in terms of electric field $\vec{\mathbf{E}}$ and the magnetic field $\vec{\mathbf{B}}$ is used. The Poynting vector is averaged value. An averaging of the monochromatic wave over the time gives the expression:

$$\left\langle \vec{\mathbf{S}} \right\rangle = \frac{\left\langle \vec{\mathbf{E}} \times \vec{\mathbf{H}} \right\rangle}{2}$$

and an averaging result over the spacial coordinates depends on the field dependencies and different for different modes.

A vector definition and approach for the energy flux definition will be used later for the normalisation of the modes. Any field decomposition over the basis functions is based on the orthogonality of the basis functions. The total field flux is a superposition of the each modal flux contribution. Regarding the Poynting vector, the averaging over the period of the structure gives that $\left\langle \vec{\mathbf{E}}_\mathbf{i} \times \vec{\mathbf{H}}_\mathbf{j} \right\rangle = 0$ for the basis functions $f_i(x)$ and $f_j(x)$ when $i \neq j$.

The Poynting vector defines the energy flux and can not be used for the evanescent waves, because it is proportional to the wave vector projection on the interesting direction:

$$\langle S_z \rangle = \mathrm{Re}\left(\frac{\left\langle E_x H_y^* - E_y H_x^* \right\rangle}{2} \right) = \mathrm{Re}\left(k_z A^2 C \left\langle f_i(x) f_i^*(x) \right\rangle \right)$$

where A is the amplitude of the wave, C is some constant depending on the base function, and $\langle f_i(x) f_i^*(x) \rangle$ is the base function averaging over the spacial coordinates (x is this case). We use the expression similar to the Poynting vector but allowing us to deal with the evanescent waves:

$$\langle P_z \rangle = \frac{\left\langle E_x \overline{H_y}^* - E_y \overline{H_x}^* \right\rangle}{2}$$

where $\overline{H}$ is the magnetic field component of the conjugated problem and for the normalisation we can write:

$$A = \sqrt{\frac{\langle P_z \rangle}{k_z C \left\langle f_i(x) \overline{f_i}^*(x) \right\rangle}}$$

2.2 The mathematical form of the solutions representation

We fix the incidence angle of the beam on the grating and denote polar angle θ and azimuthal angle ϕ. According to the diffraction theory, the reflected and the transmitted wave vector projections will be defined by the incident angles θ, ϕ, the wavelength λ of the incident light and the grating period Λ_x:

$$k_m^x = \frac{2\pi}{\lambda}\sin(\theta)\cos(\phi) + m \cdot \frac{2\pi}{\Lambda_x}$$

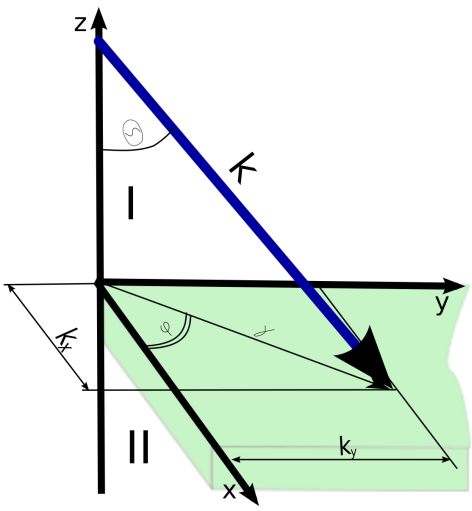


Figure 2.3: Incidence definition

the coefficient n is the order number (on simply order). Amplitudes of each wave with wave-vector k_n^x can be written in a form of the amplitudes vector $(A_0 \dots A_n)^T$.

If the incident wave has the wave-vector projection equal

$$k_{0'}^x = \frac{2\pi}{\lambda}\sin(\theta)\cos(\phi) + m' \cdot \frac{2\pi}{\Lambda_x}$$

then all the diffracted waves will have the wave-vector projections $k_{n'}^x$ equal:

$$k_{n'}^x = k_{0'}^x + n' \cdot \frac{2\pi}{\Lambda_x}$$

We can notice that solutions obtained for both these problems have the same basis in the Rayleigh orders domain.

An incident radiation thus far can be expressed as a vector of incoming beams with non-zero amplitudes for the actually radiating light.

Diffracted and reflected waves are expressed in the same basis. Thus far it was possible to build a solution in the matrix form, where amplitudes in one basis correspond to amplitudes in another basis. This representation does not increase the complexity of the problem because it is necessary to calculate amplitudes of all the outgoing waves even in the case when one

interests only several (or single) diffraction order efficiencies.

It is suitable to use complex values for the description of the amplitudes. The complex values allow to store the magnitude of the particular order along with the phase difference between each of them.

Four groups of the amplitudes can be distinguished : attributed to each side of the interface and according to the direction of propagation. Here the term interface is used intentionally to emphasise that at any cross-section of the geometry such groups can be found.

We can establish relations between these sets in different forms: a correspondence of the amplitudes (of the waves going up and down) at one side of the boundary with the amplitudes on the other side and a correspondence between the amplitudes of the waves going to the boundary (from the both sides) to the sets of the amplitudes going from the boundary. Each correspondence can be expressed in the matrix form. The matrix corresponding to the relation of the first type is called transmission T matrix and the second one is the scattering S matrix. These matrices will be described in details below.

Such correspondence can be written not only for the orders but for any two basis. So, matrices will give information of the correspondence of the modes to orders, orders to orders of modes to modes.

L.Li proposed in [20] to use R-matrix formalism providing comparison with another formalism and giving examples when using of the R-matrix approach is more suitable. This method is not widely used nowadays in the diffraction theory and we will not consider this formalism here.

2.2.1 Transition matrix

The transition matrix gives correspondence of the amplitudes (written in some basis) at one side of the boundary to the amplitudes at another side of the boundary (possibly expressed in another basis)(see fig.2.4):

$$\begin{pmatrix} b^+ \\ b^- \end{pmatrix} = T \begin{pmatrix} a^+ \\ a^- \end{pmatrix}$$

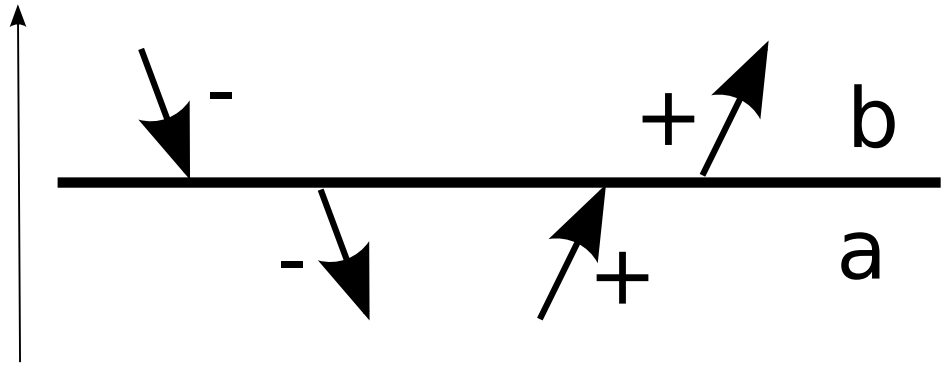


Figure 2.4: Notation of the sets at the arbitrary interface

and in opposite direction:

$$\begin{pmatrix} a^+ \\ a^- \end{pmatrix} = \overline{T} \begin{pmatrix} b^+ \\ b^- \end{pmatrix}$$

Here the vectors at the sides a and b of the boundary consist of the sub-vectors corresponding to the going upwards (superscript $+$) and downwards (superscript $-$) sets of amplitudes.In case of the infinite matrices the following equation is true:

$$\overline{T} = T^{-1}$$

This expression fails for the truncated matrices (which are used in calculations).

The sets are not obliged to be in the same basis, so sets $a^{\pm}$ and $b^{\pm}$ are, generally, in different basis. The matrix T can be divided into the sub-matrices:

$$T = \begin{pmatrix} T^{++} & T^{+-} \\ T^{+-} & T^{--} \end{pmatrix}$$

where each sub-matrix T^{ft} describes contribution of the amplitudes set with sign f to the set of amplitudes with sign t. Such subdivision of the matrix allows to see mutual influence of the amplitudes on each other. Here the terms "influence" and "correspondence" are used to stress that formulae express the equalities rather then equations. This type of matrices are useful "to get inside" the grating, for example: when amplitudes of all the orders

above the boundary are known, by multiplication of the known amplitudes by the transition matrix we obtain amplitudes of the modes inside the grating (on the other side of the boundary).

Calculation of the transition matrix of the layered structure is reduced to the multiplication of the T-matrices obtained for each elementary layer of the structure. The T-matrices multiplication is the same as for traditional matrices (subdivision is not obligatory). This property gives simplification for analytical derivations and an analytical investigation of the properties of the structure or the method. Numerical calculation demands truncation of analytically infinite matrices to some finite ones. L.Li in [21] justified possibility of such truncations and expressed the rules for the applicability of such truncation for Fourier methods.

The T-matrices application for the complex profiles is limited in numerical application due to the numerical instabilities [22]. The scattering matrices are more accurate in this regard.

2.2.2 Scattering matrix

The scattering matrix shows a correspondence between waves incident on the interface with waves leaving the interface. As a boundary here can be understood an interface between grating and the semi-infinite media or the whole grating or even stack of layers. In the terms of the fig.2.4 the scattering matrix is defined:

$$\begin{pmatrix} a^- \\ b^+ \end{pmatrix} = S \begin{pmatrix} a^+ \\ b^- \end{pmatrix}$$

The scattering matrix can be subdivided into the 2×2 matrix:

$$S = \begin{pmatrix} S_{aa} & S_{ba} \\ S_{ab} & S_{bb} \end{pmatrix} = \begin{pmatrix} S^{00} & S^{01} \\ S^{10} & S^{11} \end{pmatrix}$$

Where each sub-matrix S_{ft} gives expressions for the influence of the set at the side f to the side at the side t of the boundary. The diagonal elements of the whole S matrix are reflection coefficients of the structure. They express amplitudes of the waves on the same side of the boundary generated by the

incidence from the same side.

Depending on the implied properties of the grating and field representation basis, special properties of the submatrices and their relations can be established. These properties will not be used in this work and not covered here for this reason.

The scattering matrix can be calculated from the transition matrix of a boundary. It is also possible to combine two transition matrices written for each of the boundaries of the grating into one scattering matrix of the whole structure. This procedure can be found in [23].

The scattering matrix of the two structures with scattering matrices S_1 and S_2 can be found according to the generalised multiplication[2]:

$$S_{12} = S_1 \circ S_2 = \begin{pmatrix} S_1^{00} + S_1^{10}\left(E - S_2^{00}S_1^{11}\right)^{-1} S_2^{00} S_1^{01} & S_1^{10}\left(E - S_2^{00}S_1^{11}\right)^{-1} S_2^{10} \\ S_2^{01}\left(\left(S_1^{11}\right)^{-1} - S_2^{00}\right)^{-1}\left(S_1^{11}\right)^{-1} S_1^{01} & S_2^{11} + S_2^{01}\left(\left(S_1^{11}\right)^{-1} - S_2^{00}\right)^{-1} S_2^{10} \end{pmatrix} \tag{2.3}$$

where in each matrix under-script stands for matrix number and superscript denotes quoter of each matrix. The scattering matrices should be defined in the same basis to have physical meaning.

For the scattering matrices the multiplication is associative:

$$S_{123} = S_1 \circ S_2 \circ S_3 = S_1 \circ (S_2 \circ S_3) = (S_1 \circ S_2) \circ S_3$$

but not commutative:

$$S_{12} = S_1 \circ S_2 \neq S_2 \circ S_1 = S_{21}$$

A shifting procedure can be described for the scattering matrix in the order basis. This operation provides a scattering matrix of the layer shifted by some value x along the axis with respect to the original one. The calculation

[2]This operation is also reffered as *generalised addition*. I prefer to stick to multiplication definition because it is the first and basic operation which is introduced for groups in *Group Theory*. Scattering matrices form a group and group theory can be applied to elements of this group.

of the shifted scattering matrix is reduced to the the multiplication of the each order term by the corresponding phase shift.

2.2.3 Conversion from T matrices to S matrix

We can write T and S matrices for a certain interface. These matrices can be converted to each other according to the equations:

$$S = \begin{pmatrix} (T^{--})^{-1} T^{+-} & (T^{--})^{-1} \\ T^{++} + T^{-+} (T^{--})^{-1} T^{+-} & T^{-+} (T^{--})^{-1} \end{pmatrix}$$

where notation for T matrix is taken from the 2.2.1.

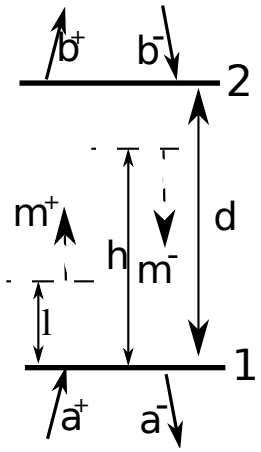


Figure 2.5: Notation used for the T matrices to S matrix conversion

We can write transition matrix for each interface of a layer. To obtain a scattering matrix of the layer there are several ways. First way is to consequently multiply the transition matrices of the boundary, layer and second boundary and convert obtained transition matrix into the scattering matrix of the layer. This approach is good for the theoretical investigation, but it leads to numerical instabilities in calculations.

Another approach is to combine transition matrices into special form and reduce calculation to matrix inversion. Suppose that we have transition matrices T_1 and T_2 for each of the interfaces. The modal representation inside the layer is defined and propagation constant of each mode is known. Suppose that modes going upwards are defined at some point l and going downwards are defined at some point h (see fig. 2.5). We can write for boundary 2:

$$\begin{pmatrix} b^+ \\ b^- \end{pmatrix} = T_2 \times U_2 \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

where U_2 is diagonal matrix, so that each term corresponds to the phase of the mode:

$$U_2 = \begin{pmatrix} \exp(jk_i^m(d-l)) & 0 \\ 0 & \exp(-jk_i^m(d-h)) \end{pmatrix}$$

for the boundary 1 we have:

$$\begin{pmatrix} a^+ \\ a^- \end{pmatrix} = T_1 \times U_1 \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

where the U_1 is matrix of the modal phases for the 1-st boundary:

$$U_1 = \begin{pmatrix} \exp(jk_i^m(0-l)) & 0 \\ 0 & \exp(-jk_i^m(0-h)) \end{pmatrix}$$

The transition matrices T_i are propagation matrices U_i can be multiplied directly. We define multiplication matrix $W_i = T_i \times U_i$ and rewrite expression for each amplitudes set:

$$\begin{pmatrix} b^+ \\ 0 \end{pmatrix} = \begin{pmatrix} W_2^{++} & W_2^{-+} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ a^- \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ W_1^{+-} & W_1^{--} \end{pmatrix} \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ b^- \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ W_2^{+-} & W_2^{--} \end{pmatrix} \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

$$\begin{pmatrix} a^+ \\ 0 \end{pmatrix} = \begin{pmatrix} W_1^{++} & W_1^{-+} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

Regrouping expression for b^+, a^- and b^-, a^+ we get:

$$\begin{pmatrix} a^- \\ b^+ \end{pmatrix} = \begin{pmatrix} W_1^{+-} & W_1^{--} \\ W_2^{++} & W_2^{-+} \end{pmatrix} \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

$$\begin{pmatrix} a^+ \\ b^- \end{pmatrix} = \begin{pmatrix} W_1^{++} & W_1^{-+} \\ W_2^{+-} & W_2^{--} \end{pmatrix} \begin{pmatrix} m^+ \\ m^- \end{pmatrix}$$

From these expression follows the S matrix expression:

$$S = \begin{pmatrix} W_1^{+-} & W_1^{--} \\ W_2^{++} & W_2^{-+} \end{pmatrix} \begin{pmatrix} W_1^{++} & W_1^{-+} \\ W_2^{+-} & W_2^{--} \end{pmatrix}^{-1}$$

This way to obtain scattering matrix of the layer with depth d is more stable in numerical sense.

There is similar expression for the opposite transition matrices $\overline{T}$ but we will skip derivation of the expression here.

2.3 Slicing

The techniques developed for the vertical walls can not be directly applied to arbitrary profile. Meanwhile, there are profiles, a direct calculation of which can not be done with any existing method. It is possible to approximate the original profile as a stack of gratings that are convenient for calculating the specific method. In the case of modal methods, such approximation is the representation of arbitrary profile in the form of stairs (stair-like approximation). The validity of such approximation can be found in [24]. For each slice the S-matrix is calculated. After the matrices for each layer are calculated one obtains the final matrix by consequent multiplication of the scattering matrices according to the (2.3). Generally, the representation of the full profile in the form of a stack of layers is called slicing. It seems to be clear that if more layers are used in the representation then the closer the result will be to the original profile. This method increases time of the complete structure calculation linearly with the number of layers used for representation of the profile. In addition to this limitation, it should be kept in mind that S-matrix for each layer is calculated with a definite error and that in a matrix multiplication the error accumulates. As a consequence, there is an implicit upper limit of the number of layers used. Thus, in order to achieve the best slicing results, the convergence of interesting values should be investigated as a function of the layers number. Modal methods shows the decrease of the error if the number of variables used in the simulation increases. Taking

into account both the dependencies we find that for the slicing technique the problem to obtain the most accurate solution reduces to the optimisation on the two parameters problem, i.e. to the search for the optimum over the two-dimensional field.

2.3.1 Slicing technique as finite summation and integration limit

The multiplication of the S matrices, written for the slices of some depth dh, gives the total scattering matrix. The slicing technique does not imply any restriction on the number of slices and depth of each slice.

The final scattering matrix is a function of number of slices and converges to final result. In case of slicing we are multiplying S-matrix, each depending on the profile and the depth of the slice. The numerical integration utilises summation of the products of the function value calculated at the point multiplied with a step corresponding to this point. Numerical integration and slicing are somehow similar in terms of the calculation of the values at some coordinate and consequent multiplication (addition) of the obtained results. There is an optimum for the number of layers, which reaches a minimum error and provides maximum reliability. Also similarity allows to use existing methods of numerical integration for calculating of the scattering matrix of complex structures.

2.4 Methods to validate results

After the method is implemented and the first results are obtained a question appears whether the results are valid. The first solution is to obtain the results with another method and to compare the results by the two methods. This approach to the method validation works only if another method exists and provides reference results for the particular grating profile. Even when such method exists and results are comparable the following question arises: what results are more accurate? to what results can we trust more?

A first and evident solution lies in check of the energy conservation. This

can be applied only for the lossless materials but also provide some estimation for the lossy materials.

According to the energy conservation theorem the sum of the energy efficiencies (squared amplitudes of the outgoing orders {divided by permittivity in case of the magnetic field amplitudes}) should be equal to unity. It is generally called the *energy balance criterion*. The physical interpretation is simple: the incident energy is equal to the diffracted energy. This method is not applicable for some methods which are based on this conservation or utilize somehow this property (C-method for example).

Another validation criterion comes from the symmetry of the solution with regard to the substitution of the incoming beams with the outgoing and vice-verse. This validation called *the reciprocity theorem* [25].

These methods are the first benchmarks for those who develops a new method or improve existing ones.

2.5 Method benchmark

If the developed method satisfies criteria described above another treatment of the method can should be applied. The method should be stable with regard to the slight variation of the initial parameters. The resulting solutions should be close to each other (except special cases like resonances). Varying the parameters the developer can check if the method stable or to define for which domain the developed method is suitable.

The numerical methods has (as a rule) some parameter, characterising only the certain run of the method: number of orders under consideration, number of points considered or some step parameter. Another validity check is based on the convergence of the results to some point with the increase of that specific for the simulation parameter. It is also evident that such convergence exists up to some limit when accumulated errors will surpass benefits from more accurate problem representation.

We can consider the error of the result as a function of the inverse of characterising parameter $\frac{1}{m}$ (for example number of modes). The analysis provides accurate results, so we can write that for infinite number of modes

we get zero error:

$$e(\frac{1}{\infty}) = V_t - V(\frac{1}{\infty}) = 0$$

where V_t is the true value and the $V(\frac{1}{m})$ is the value obtained with finite number of modes m. For any finite number of modes there will always be an error. When we can not find real value we can use self-convergence to see if the error decreases with the increasing number of modes.

Results convergence is also a validation method and can be used along with other validation techniques.

2.6 The processor precision. The floating point precision. The numerical calculations

The performance of the computers grows quite rapidly. Along with the performance the available for computation memory also increases. Both these facts gives possibility to implement methods which were impossible to implement in the past. High performance of the processor also allows to concentrate on the ideas to be implemented and pass more and more computational work to the processor. This way has certain drawbacks. First of all, there is the productivity reduction. The second reason is the error accumulation. Each floating point value is presented with the finite precision. The processor precision α is a value so that

$$A \cdot (1 + \alpha) = 1 \cdot A$$

This equation should be treated that there always exists some value, so that adding this value to one will not influence on the result (value still be one). This value depends on the mantissa size, on the number of bits used for the fractional part representation. For example, a single precision type has accuracy equals to $5.960e-08$ and a double precision type has approximately 16 decimal digits.

If appears that better accuracy is necessary for the calculations, the special techniques should be used [26]. That will reduce computation perfor-

mance and the trade-off should be estimated before applying these methods.

Finite summations can not be avoided. And it should be kept in mind that each floating point value has already representation error (stored value differs from the calculated one according to the rounding rules). These values are not equally distributed in sense of the rounding up or down and summation can significantly influence the final result.

This topic is not widely discussed in physical society but in the computer science there can be found some explanations and more information [27].

2.7 Fourier transformation

The key instrument for he RCWA and Chandezon's methods is the Fourier transform. The Fourier transform is a mathematical operation that decomposes a signal into its constituent frequencies. The original signal depends on space (or time), and therefore is called the space (or time) domain representation of the signal, whereas the Fourier transform depends on frequency and is called the frequency domain representation of the signal. The term Fourier transform refers both to the frequency domain representation of the signal and the process that transforms the signal to its frequency domain representation. [3]

In mathematical terms, the Fourier transform 'transforms' one complex-valued function of a real variable into another. In effect, the Fourier transform decomposes a function into oscillatory functions. The Fourier transform and its generalisations are the subject of Fourier analysis. In this specific case, both the time and frequency domains are unbounded linear continua. It is possible to define the Fourier transform of a function of several variables, which is important for instance in the physical study of wave motion and optics. It is also possible to generalise the Fourier transform on discrete structures such as finite groups. The efficient computation of such structures, by fast Fourier transform, is essential for high-speed computing.

Definition There are several common conventions for defining the Fourier

[3] This section is partial copy of the article, which can be found on the [28] and further links.

transform $\hat{f}$ of an integrable function $f : \mathbb{R} \to \mathbb{C}$. Here we will use the definition:

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\ e^{-2\pi i x \xi}\, dx,$$

for every real number ξ.

When the independent variable x represents time, the transform variable ξ represents frequency (in Hertz). Under suitable conditions, f can be reconstructed from $\hat{f}$: by the inverse transform:

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)\ e^{2\pi i x \xi}\, d\xi$$

for every real number x.

The Fourier transform on Euclidean space is treated separately, in which the variable x often represents position and ξ momentum.

There is a close connection between the definition of Fourier series and the Fourier transform for functions f which are zero outside of an interval. For such a function we can calculate its Fourier series on any interval that includes the interval where f is not identically zero. The Fourier transform is also defined for such a function. As we increase the length of the interval on which we calculate the Fourier series, then the Fourier series coefficients begin to look like the Fourier transform and the sum of the Fourier series of f begins to look like the inverse Fourier transform. To explain this more precisely, suppose that T is large enough so that the interval $[-T/2, T/2]$ contains the interval on which f is not identically zero. Then the n-th series coefficient c_n is given by:

$$c_n = \int_{-T/2}^{T/2} f(x) \exp(-2\pi i (n/T) x) dx$$

Comparing this to the definition of the Fourier transform it follows that $c_n = \hat{f}(n/T)$ since $f(x)$ is zero outside $[-T/2, T/2]$. Thus the Fourier coefficients are just the values of the Fourier transform sampled on a grid of width $1/T$. As T increases the Fourier coefficients more closely represent the Fourier transform of the function.

Under appropriate conditions the sum of the Fourier series of f will equal the function f. In other words f can be written:

$$f(x) = \frac{1}{T} \sum_{n=-\infty}^{\infty} \hat{f}(n/T)\ e^{2\pi i (n/T)x} = \sum_{n=-\infty}^{\infty} \hat{f}(\xi_n)\ e^{2\pi i \xi_n x} \Delta\xi$$

where the last sum is simply the first sum rewritten using the definitions $\xi_n = n/T$, and $\Delta\xi = (n+1)/T - n/T = 1/T$.

In the study of Fourier series the numbers c_n could be thought of as the "amount" of the wave in the Fourier series of f. Similarly, as seen above, the Fourier transform can be thought of as a function that measures how much of each individual frequency is present in our function f, and we can recombine these waves by using an integral (or "continuous sum") to reproduce the original function.

The field decomposition represented as the Rayleigh series can be treated as the Fourier decomposition of the total electromagnetic field and application of the Fourier technique seems to be very suitable for the diffraction problem resolution.

Very important property, which defines application of the Fourier decomposition in the differential equation system solutions, is the following:

$$\alpha f(x) + \hat{\beta} \cdot g(x) = \alpha \hat{f(x)} + \beta \hat{g(x)}$$

$$\left(\frac{\hat{f(x)}}{\partial x} \right) = i\omega \hat{f(x)}$$

For the convolution function $h(x)$ of functions $f(x)$ and $g(x)$ defined like $h(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(y) g(x-y)\, dy$ Fourier transform is

$$\hat{h}(\xi) = \hat{f}(\xi) \cdot \hat{g}(\xi)$$

the dual property for the $h(x) = f(x) \cdot g(x)$ is

$$\hat{h}(\xi) = \hat{f}(\xi) * \hat{g}(\xi)$$

. According to these formulae the differential operators in Maxwell's equa-

tions are reduced to the multiplication.

Another property of the Fourier decompositions concerns theirs multiplication. Distributions can be multiplied only when there are no common singular points. This rule lead to several publications showing convergence improvement of the obtained results and to the formulation of the Fourier method for the gratings with slanted walls.

2.7.1 Fast Fourier transform

A fast Fourier transform (FFT) is an efficient algorithm to compute the discrete Fourier transform (DFT) and its inverse. There are many distinct FFT algorithms involving a wide range of mathematics, from simple complex-number arithmetic to group theory and number theory; this section gives an overview of the available techniques and some of their general properties, while the specific algorithms are described in subsidiary articles linked below. [4]

A DFT decomposes a sequence of values into components of different frequencies. This operation is useful in many fields (see discrete Fourier transform for properties and applications of the transform) but computing it directly from the definition is often too slow to be practical. An FFT is a way to compute the same result more quickly: computing a DFT of N points in the naive way, using the definition, takes O(N2) arithmetical operations, while an FFT can compute the same result in only O(N log N) operations. The difference in speed can be substantial, especially for long data sets where N may be in the thousands or millions—in practise, the computation time can be reduced by several orders of magnitude in such cases, and the improvement is roughly proportional to N / log(N). This huge improvement made many DFT-based algorithms practical; FFTs are of great importance to a wide variety of applications, from digital signal processing and solving partial differential equations to algorithms for quick multiplication of large integers.

[4] This section is partial copy of the article, which can be found on the [29] and further links.

The most well known FFT algorithms depend upon the factorisation of N, but (contrary to popular misconception) there are FFTs with O(N log N) complexity for all N, even for prime N. Many FFT algorithms only depend on the fact that $\exp(-\frac{2\pi i}{N})$ is an $N-$th primitive root of unity, and thus can be applied to analogous transforms over any finite field, such as number-theoretic transforms.

Since the inverse DFT is the same as the DFT, but with the opposite sign in the exponent and a 1/N factor, any FFT algorithm can easily be adapted for it.

2.8 Existing methods developed for one-dimensional gratings

The problem to obtain exact solution for the diffraction problem on the one-dimensional grating has a long history. A high result accuracy and solution tolerance to the grating parameters are dictated by the applications of the elements: mirrors in lasers, sensors, filters. The accuracy of the model provides an answer if it is possible to apply solutions in the desired domain and to produce it with a present technology.

Starting from these assumptions, the most important criteria become the accuracy and reliability of the results provided with a particular method. Another important factor, when choosing a method to be used, is a calculation time and a memory consumption. This factor plays important role when scanning over parameters of the grating (period, depth, profile, materials) or wavelengths is performed [30]. When the slicing technique is applied, the method's performance also plays an important role.

It is possible to make a correspondence between grating profile and more suitable method for resolution of the diffraction problem on the grating with such profile. This correspondence arises from the representation of the profile and resolution method.

2.8.1 C-method

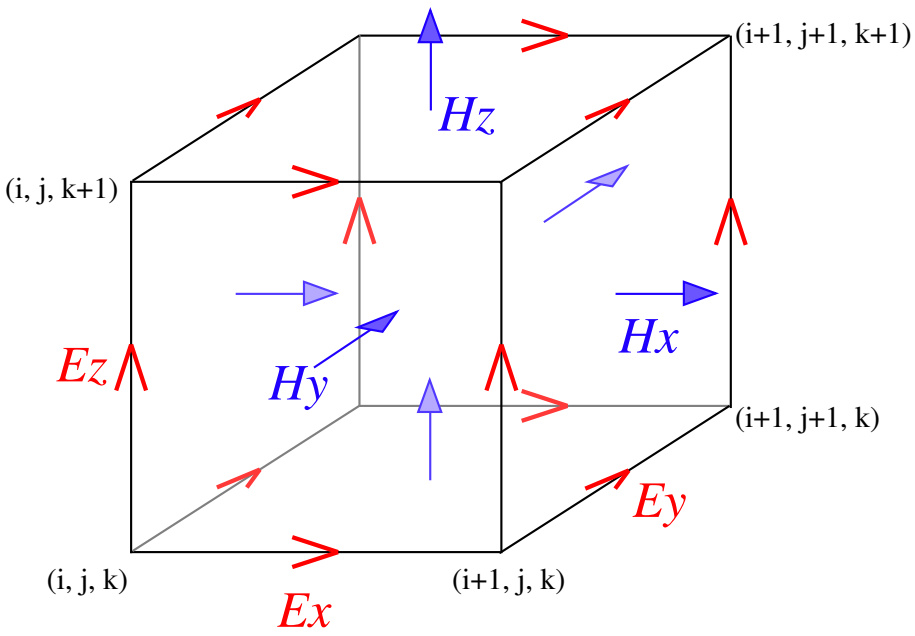
The C-method was called after the author of this approach. Chandezon et al. introduced this approach in the [31]. In this paper the authors showed a method to calculate diffraction efficiencies for both polarisation of the incoming beam.

The core idea of the method is to write the Maxwell's equations in the "translation coordinate system". This system is bound with one of the surfaces of the grating, so that corrugated surface of the grating becomes a coordinate plane in the new system.

The Maxwell's equations are then reformulated in terms of the new coordinate system in the Fourier space. This action leads to simplification of the boundary condition and simplified problem resolution.

As the new coordinate system is bound with the profile then the profile shape should be characterised by the invertible function. This method is applicable for the smooth profiles or at least for the profiles without overhanging.

2.8.2 FDTD method



Finite-difference time-domain [5] (FDTD) is a popular computational electrodynamics modelling technique. It is considered easy to understand and easy to implement in software. Since it is a time-domain method, solutions can cover a wide frequency range with a single simulation run.

The FDTD method belongs in the general class of grid-based differential time-domain numerical modelling methods. The time-dependent Maxwell's equations (in partial differential form) are discretised using central-difference approximations to the space and time partial derivatives. The resulting finite-difference equations are solved in either software or hardware in a leapfrog manner: the electric field vector components in a volume of space are solved at a given instant in time; then the magnetic field vector components in the same spatial volume are solved at the next instant in time; and the process is repeated over and over again until the desired transient or steady-state electromagnetic field behaviour is fully evolved.

In order to use FDTD a computational domain must be established. The computational domain is simply the physical region over which the simulation will be performed. The E and H fields are determined at every point in space within that computational domain. The material of each cell within the computational domain must be specified. Typically, the material is either free-space (air), metal, or dielectric. Any material can be used as long as the permeability, permittivity, and conductivity are specified.

Once the computational domain and the grid materials are established, a source is specified. The source can be an impinging plane wave, a current on a wire, or an applied electric field, depending on the application

Since the E and H fields are determined directly, the output of the simulation is usually the E or H field at a point or a series of points within the computational domain. The simulation evolves the E and H fields forward in time.

Processing may be done on the E and H fields returned by the simulation. Data processing may also occur while the simulation is ongoing.

While the FDTD technique computes electromagnetic fields within a com-

[5] This section is partial copy of the article, which can be found on the [32] and further links.

pact spatial region, scattered and/or radiated far fields can be obtained via near-to-far-field transformations[33].

The basic FDTD space grid and time-stepping algorithm trace back to a seminal 1966 paper by Kane Yee in IEEE Transactions on Antennas and Propagation [34]. The descriptor "Finite-difference time-domain" and its corresponding "FDTD" acronym were originated by Allen Taflove in a 1980 paper in IEEE Transactions on Electromagnetic Compatibility [35].

Since about 1990, FDTD techniques have emerged as primary means to computationally model many scientific and engineering problems dealing with electromagnetic wave interactions with material structures. Current FDTD modelling applications range from near-DC (ultralow-frequency geophysics involving the entire Earth-ionosphere waveguide) through microwaves (radar signature technology, antennas, wireless communications devices, digital interconnects, biomedical imaging/treatment) to visible light (photonic crystals, nanoplasmonics, solitons, and biophotonics)[36]. In 2006, an estimated 2,000 FDTD-related publications appeared in the science and engineering literature.

2.8.3 Rayleigh method

Lord Rayleigh in the 1907 put forward the diffraction theory by a reflection grating, where he assumed the discrete set of upward reflected spectral waves (together with the incident field) to be a complete description of the total field up to the boundary surface of the grating. In the [37] *van den Berg* and *Fokkema* have shown in a fairly simple way the conditions under which the Rayleigh hypothesis holds rigorous for wide class of profiles. These authors applied later [38] Rayleigh hypothesis to the perturbation of the plane surface. Recently the validity of the Rayleigh hypothesis was demonstrated for the deep gratings [39, 40]. Thus far this method is valid and can be applied for the diffraction efficiency computation among the other methods.

2.8.4 Rigorous Coupled Wave Analysis (Fourier-modal method)

Another method, which uses Fourier transform, was introduced by Knop [3] and later adopted by Moharam and Gaylord in [41]. This method represents the field inside the grating as a superposition of Fourier harmonics. The plane-wave basis allows representation in the Fourier space where differential operators are represented as a multiplication of the field components by the corresponding frequencies.

This method is quite easy to understand and implement. This quality gave a popularity and a wide application to this method. The most suitable profile for this method is lamellar grating. It can be binary grating or containing more then two components over a period.

The convergence of this method for the TM polarisation was the question to solve for a long time. The explanation of the slow convergence and a way to get faster convergence was given in the Ph.Lalanne's paper [42]. His explanation was later reformulated by L.Li in [43].

For the arbitrary profiles the slicing technique is used when each of the slices is represented as a lamellar grating. Later in the [44], after the Néviere Popov's paper [17], the method was extended to be applicable for the arbitrary profiles sliced into the trapezoidal profiles.

Previously described methods provide far-field efficiencies (except FDTD, which provides a near field distribution at the grating surface), they answer what will be the reflected or transmitted light but do not allow "to look inside" the grating.

Not all the existing methods are covered in the section above. There are also intergal methods, generalized sources method, fictional sources method, and others. Another methods are not covered here because they use ideas and approaches totally different from those which are used in modal methods.

The important information about the electromagnetic solutions inside the grating (like field distribution, energy flux) are not described precisely by these methods and it is necessary to apply other techniques or to use true-modal method.

2.8.5 True-modal method

The true-modal method is based on the representation of the field inside the grating as a superposition of the modes, the basic solutions of the particular problem consisting of the boundary conditions and incidence conditions.

Each mode is characterised with some propagation constant (propagation speed) and field distribution. The solution of the problem attribute the amplitudes to each mode to fulfil the Maxwell's equations. Modes are solutions of the electromagnetic equations and the real light propagation inside the grating is describe with this method in the most accurate fashion.

This method allows to look inside the grating, to investigate the processes of the energy propagation, inter-modal exchange and explain some effects like wide-band reflection or polarisation. The paper [23] gives good example of the explanation provided with the true-modal method.

The methods can be separated to the calculation methods, which utilise some mathematical tricks to solve Maxwell's equations, and the models. As a model here I understand some problem description and development of the model based on the physical processes. From this point of view, only true-modal method is a model based on the physical process of the light propagation inside the grating.

Chapter 3

Building a true-modal solution for a one-dimensional grating

The one-dimensional method is well adopted to be used as a base method for constructing the two-dimensional basis of a simple structure used later for investigation. It was also shown that true-modal method can be used for the metal gratings under conical incidence[45]. This chapter revisit the field of the true-modal method for one-dimensional binary grating structures. Special attention is given to the fact that analytical solutions are not limited to the wave-vector projection been real and applicable for the imaging complex values.

3.1 Introduction

The true-modal technique will be described in detail. Special attention is devoted to this method because it will be intensively used in the following development of the true-modal method for two-dimensional structures. The first section is devoted to describing the geometry of the problem, and introducing the basic variables and parameters. Then to reduce the difficulty of the problem, all the solutions are partitioned into two main groups. Such partition is based on the polarisation definition. Dispersion equations will be provided for each type of the polarisation. These equations define a set of

propagation-constant values for binary gratings. A generalisation for grating structures composed of more then two components is provided after that. A short explanation of the conjugated problem will be introduced to get an orthogonal basis, which allows to normalise mode amplitude distributions along the period. Unique amplitude definition, which is based on the orthogonality, allows to compare modes with each other in terms of the energy transition. Formulae for the overlap integrals will then be given, providing an explanation based on an understandable physical insight. The last steps for the resolution of the one-dimensional problem are to construct and calculate the transition matrix components (which are based on the overlap integral values) and the final calculation of the scattering matrix based on the two transition matrices.

3.2 Description of one-dimensional grating

From a technological point of view, gratings consisting of only two materials are the most convenient (and simple) in production. The profile functions can be sinusoidal or lamellar (or close to one of them) . Technological limitations sometimes lead to gratings with a trapezoidal profile function (when walls are slightly slanted towards the top of the grating). In case of small deviations it is still possible to consider approximate solutions as the walls were vertical.

A typical problem geometry is depicted in fig.3.1. A coordinate system is connected with the grating in the following way: the $x-$axis coincides with the direction of the grating periodicity; z direction is perpendicular to the grating surface (positive direction is taken so that reflected waves have positive projections) and $y-$axis is taken, so that xyz coordinate system is right handed. A plane wave with wavelength λ and wave vector $k = \frac{2\pi}{\lambda}$ illuminates the grating at polar angle Θ and azimuthal angle ϕ (taken from the x direction anticlockwise). Notation of incident wave vector projections are the same as in sec. 2.1.1. We want to construct Maxwell's equations solution for the grating region, so that the dependence on the $z-$coordinate will be separated and each function, describing the field distribution inside

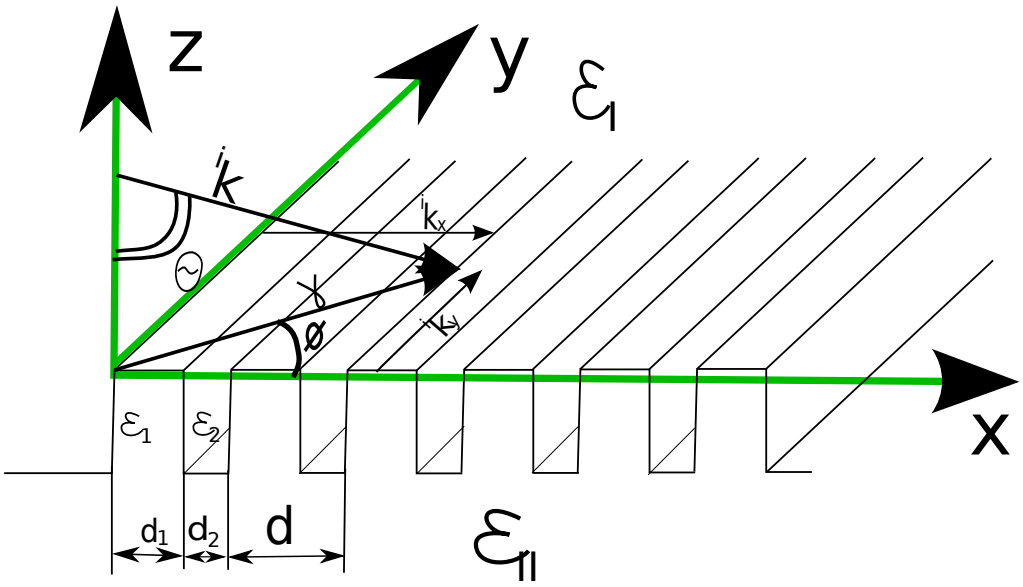


Figure 3.1: Typical binary grating configuration

the grating, will have the following view:

$$f(x, y, z) = g(x, y) \cdot h(z)$$

There are many functions that satisfy Maxwell's equations, and admitting the separation of the depth variable. One can select a group of functions among these functions such that any distribution of the field in the grating is represented as the composition of these functions. In other words, it is possible to construct a functional basis so that any solution inside the grating would be represented as a basis functions composition with amplitudes defined from the problem formulation.

This approach, when a general solution is decomposed into functional (modal) basis, is called modal method. It is initially claimed, thus far, that all solutions will be represented by the modes (basic functions). In the one-dimensional grating case with walls perpendicular to the grating surface, these modes are three dimensional functions such that the coordinate dependence along the coordinate with the translational symmetry ($y-$ axis) is dictated only by incident wave vector projection, the depth-dependence is defined by the propagation constant of the mode (and wave vector projection

along the translational symmetry) and the coordinate dependence along the periodicity direction is represented as a function with predefined properties. In geometry of the fig.3.1, the dependence would look like:

$$f(x,y,z) = \psi(x)g(y)\exp(\pm jk^z(\beta)z)$$

where β is the modal constant (modal propagation constant), $k^z(\beta)$ is the projection of the modal constant on the depth-coordinate, $\psi(x)$ is the modal distribution function, corresponding to the particular value of the modal constant and $g(y)$ is the function along the translation symmetry direction (unique for all mode),

3.3 Mode polarisation

We can write down the coordinate dependence of the reflected and transmitted plane waves:

$$\begin{aligned} \vec{\mathbf{E}}_o(x,y,z) &= \vec{\mathbf{E}}\exp(jk_o^x x)\exp(jk^y y)\exp(jk^z z)\exp(-j\omega t) \\ \vec{\mathbf{H}}_o(x,y,z) &= \vec{\mathbf{H}}\exp(jk_o^x x)\exp(jk^y y)\exp(jk^z z)\exp(-j\omega t) \end{aligned} \quad (3.1)$$

where o is order number, $k_o^x = \overset{i}{k}^x + \frac{2\pi}{d}$ is Rayleigh's order decomposition, $k^y = \overset{i}{k}^y$, $k^z = \pm\sqrt{\omega\varepsilon_s - (k_o^x)^2 - (k^y)^2}$ (positive for reflected orders and negative for transmitted waves) and s stands for the number of semi-infinite media. It is easy to verify that this representation satisfies the Maxwell's equations (2.1), or written in derivatives:

$$\begin{aligned} -j\omega\varepsilon E_x &= \tfrac{\partial H_z}{\partial y} - \tfrac{\partial H_y}{\partial z} & j\omega\mu H_x &= \tfrac{\partial E_z}{\partial y} - \tfrac{\partial E_y}{\partial z} \\ -j\omega\varepsilon E_y &= \tfrac{\partial H_z}{\partial x} - \tfrac{\partial H_x}{\partial z} & j\omega\mu H_y &= \tfrac{\partial E_z}{\partial x} - \tfrac{\partial E_x}{\partial z} \\ -j\omega\varepsilon E_z &= \tfrac{\partial H_y}{\partial x} - \tfrac{\partial H_x}{\partial y} & j\omega\mu H_z &= \tfrac{\partial E_y}{\partial x} - \tfrac{\partial E_x}{\partial y} \end{aligned}$$

We now construct the basis functions, which will continue to represent the field inside the grating. We operate with the solutions given by the modal constant equals to β. So we have that $\beta^2 = (k^z)^2 + (k^y)^2$. We fix this dependence in exponential form $\exp(jk^y y)\exp(\pm jk^z z)$ and rewrite Maxwell's

equations:

$$\begin{aligned}
\omega\varepsilon(x)E_x(x,y,z) &= jk_zH_y(x,y,z) - jk_yH_z(x,y,z) \\
-j\omega\varepsilon(x)E_y(x,y,z) &= \tfrac{\partial H_z(x,y,z)}{\partial x} - jk_zH_x(x,y,z) \\
-j\omega\varepsilon(x)E_z(x,y,z) &= \tfrac{\partial H_y(x,y,z)}{\partial x} - jk_yH_x(x,y,z)
\end{aligned}$$

$$\begin{aligned}
\omega\mu(x)H_x(x,y,z) &= jk_yE_z(x,y,z) - jk_zE_y(x,y,z) \\
j\omega\mu(x)H_y(x,y,z) &= \tfrac{\partial E_z(x,y,z)}{\partial x} - jk_zE_x(x,y,z) \\
j\omega\mu(x)H_z &= \tfrac{\partial E_y(x,y,z)}{\partial x} - jk_yE_x(x,y,z)
\end{aligned}$$

Observing the equation system written in that form it can be seen, that using $H_x(x,y,z) = 0$ for the first three equations leads to the field description as a one-dimensional function with its amplitude vector laying in the plane of walls $(y-z)$. Same action can be applied to the second triple of the equations with $E_x = 0$. Introducing this restriction, we declare that for one type of solutions the electric vector will be lying parallel to the walls of the grating and for the other this property will be valid for the magnetic vector. In literature, waves with the electric vector laying along the grating walls called the TE (transverse electrical), and another type of wave is called TM (transverse magnetic). We will use the same definition. Introducing the amplitude of the modal field $a^{e,h}$ where the superscript denotes electrical or magnetic field amplitude, we may write down expression for the $y-$ and $z-$ projections and thereby reduce the problem to one-dimensional function. After substitution of the projection expressions into Maxwell's equations we obtain:

$$\begin{array}{rcc}
 & TM \quad mode & TE \quad mode \\
E_x = & \frac{\beta}{\omega\varepsilon(x)}a^h\psi^h(x) & 0 \\
E_y = & \frac{1}{j\omega\varepsilon(x)}\frac{k^y}{\beta}a^h\psi^{h\prime}(x) & \frac{k^z}{\beta}a^e\psi^e(x) \\
E_z = & -\frac{1}{\omega\varepsilon(x)}\frac{k^z}{\beta}a^h\psi^{h\prime}(x) & -\frac{k^y}{\beta}a^e\psi^e(x) \\
H_x = & 0 & -\frac{\beta}{\omega\mu}a^e\psi^e(x) \\
H_y = & \frac{k^z}{\beta}a^h\psi^h(x) & -\frac{1}{j\omega\mu}\frac{k^y}{\beta}a^e\psi^{e\prime}(x) \\
H_z = & -\frac{k^y}{\beta}a^h\psi^h(x) & \frac{1}{j\omega\mu}\frac{k^z}{\beta}a^e\psi^{e\prime}(x)
\end{array} \tag{3.2}$$

field expressions where $\psi^e(x)$ and $\psi^h(x)$ are of a particular TE or TM mode. In what follows derivative over x is denoted with prime ($f'(x) = \frac{\partial f}{\partial x}$).

If we express the field inside the grating with M modes of each type, then we have $2M$ amplitudes which describe all the fields projections of waves propagating upwards and $2M$ amplitudes of the modes going downwards (in the following expression k^z is supposed to be signed):

$$\begin{aligned}
E_x = & \sum_{i=0}^{M} & \frac{\beta_i^h}{\omega\varepsilon(x)} a_i^h \psi_i^h(x) \\
E_y = & \sum_{i=0}^{M} \left(\frac{k_i^{ze}}{\beta_i^e} a_i^e \psi_i^e(x) \right. & \left. + \frac{1}{j\omega\varepsilon(x)} \frac{k^y}{\beta_i^h} a_i^h \psi_i^{h\prime}(x) \right) \\
E_z = & \sum_{i=0}^{M} \left(-\frac{k^y}{\beta_i^e} a_i^e \psi_i^e(x) \right. & \left. - \frac{1}{\omega\varepsilon(x)} \frac{k_i^{zh}}{\beta_i^h} a_i^h \psi_i^{h\prime}(x) \right) \\
H_x = & \sum_{i=0}^{M} -\frac{\beta_i^e}{\omega\mu} a_i^e \psi_i^e(x) & \\
H_y = & \sum_{i=0}^{M} \left(-\frac{1}{j\omega\mu} \frac{k^y}{\beta_i^e} a_i^e \psi_i^{e\prime}(x) \right. & \left. + \frac{k_i^{zh}}{\beta_i^h} a_i^h \psi_i^h(x) \right) \\
H_z = & \sum_{i=0}^{M} \left(\frac{1}{j\omega\mu} \frac{k_i^{ze}}{\beta_i^e} a_i^e \psi_i^{e\prime}(x) \right. & \left. - \frac{k^y}{\beta_i^h} a_i^h \psi_i^h(x) \right)
\end{aligned} \tag{3.3}$$

where superscript e and h of the amplitudes, modal constants and modal functions stands for TE and TM polarisation, $k_i^{z,p} = \pm\sqrt{\beta_i^p - (k^y)^2}$ is the projection of the modal constant on the z direction and index i stands for the mode number.

3.4 Invariance of the polarisation

A special but very useful case for common knowledge comes from the irradiation of the grating with radial angle ϕ equals zero (called non-conical mount). In non conical mount is the configuration where polarisation of the incident wave, reflected and transmitted orders remains the same. In that mount k^y projection of the incident

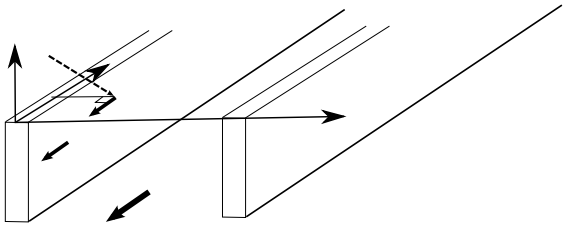


Figure 3.2: Polarisation invariance

wave is 0.

If we consider an incident wave with TE polarisation, it can be seen, that such configuration would excite only TE-modes (inside the grating region). Modes in their turn would be re-radiated into the plane waves with the same polarisation at both the interfaces of the grating.

The same derivation is right for the TM polarised plane wave and TM modes.

That finding can be used later to check the validity of the derived expressions.

3.5 The dispersion equation

Now we consider the system of equations (3.2) separately for each polarisation. Thus, for the case of TE polarisation, the system can be reduced to an equation of the form:

$$\left(\frac{\psi^e(x)'}{\omega\mu(x)}\right)' + \omega\varepsilon(x)\psi^e(x) = (\beta^e)^2\,\frac{\psi^e(x)}{\omega\mu(x)}$$

for TM polarisation, the equation becomes:

$$\left(\frac{\psi^h(x)'}{\omega\varepsilon(x)}\right)' + \omega\mu(x)\psi^h(x) = \left(\beta^h\right)^2\,\frac{\psi^h(x)}{\omega\varepsilon(x)}$$

From the equations we can see that in the case of TE modes, the functions $\psi^e(x)$ and $\psi^{e\prime}(x)$ are continuous, while for TM case, the ration $\frac{\psi^{h\prime}(x)}{\omega\varepsilon(x)}$ is continuous.

Now we construct functions that satisfy the dispersion equation for various polarisations in case of a binary grating. Within each of the parts of the period we define our function as the sum of 2 plane waves propagating towards each other and starting on opposite walls of the slab. In this case,

the function will be piecewise and represented as:

$$\psi_i(x) = \begin{cases} \psi_{1.i}(x) \\ \psi_{2.i}(x) \end{cases} = \begin{cases} f_1 \exp(jk^x_{1.i}x) + f_2 \exp(-jk^x_{1.i}x) & ,0 \le x \le d_1 \\ f_3 \exp(jk^x_{2.i}x) + f_4 \exp(-jk^x_{2.i}x) & d_1 \le x \le d_1 + d_2 \end{cases}$$

where $k^x_{n.i} = \sqrt{\omega^2 \mu \varepsilon_n - {\beta_i}^2}$. Then, from Maxwell's equations, which require continuity of the tangential projection of the fields, and under periodicity conditions imposed on the function, we find that the amplitudes of modal waves should satisfy the following criteria (for the TE mode):

$$\begin{cases} \psi_{1.i}(d_1 - 0) &= \psi_{2.i}(d_1 + 0) \\ \psi'_{1.i}(d_1 - 0) &= \psi'_{2.i}(d_1 + 0) \\ \psi_{2.i}(d - 0) &= \psi_{1.i}(d + 0) \\ \psi'_{2.i}(d - 0) &= \psi'_{1.i}(d + 0) \\ \psi_i(x) &= \psi_i(x + d)\exp(j^i k^x d) \end{cases}$$

or, after substitution of the representation of our function for a binary grating, we obtain:

$$\begin{cases} f_1 \exp(jk^x_{1.i}d_1) + f_2 \exp(-jk^x_{1.i}d_1) = f_3 \exp(jk^x_{2.i}d_1) + f_4 \exp(-jk^x_{2.i}d_1) \\ k^x_{1.i} f_1 \exp(jk^x_{1.i}d_1) - k^x_{1.i} f_2 \exp(-jk^x_{1.i}d_1) = k^x_{2.i} f_3 \exp(jk^x_{2.i}d_1) - k^x_{2.i} f_4 \exp(-jk^x_{2.i}d_1) \\ f_3 \exp(jk^x_{2.i}d) + f_4 \exp(-jk^x_{2.i}d) = f_1 \exp(jk^x_{1.i}d) + f_2 \exp(-jk^x_{1.i}d) \\ k^x_{2.i} f_3 \exp(jk^x_{2.i}d) - k^x_{2.i} f_4 \exp(-jk^x_{2.i}d) = k^x_{1.i} f_1 \exp(jk^x_{1.i}d) - k^x_{1.i} f_2 \exp(-jk^x_{1.i}d) \\ f_1 \exp(jk^x_{1.i}d) + f_2 \exp(-jk^x_{1.i}d) = [f_1 \exp(jk^x_{1.i}0) + f_2 \exp(-jk^x_{1.i}0)] \exp(j^i k^x d) \end{cases}$$

These 4 equations contain 5 unknowns: 4 amplitudes and one hidden parameter β^2. This can also be represented in a matrix form where a linear superposition of the rows corresponding to the β^2 will be equal to zero, and as a result the determinant of the matrix is equal zero. Rigorously speaking, this system has an infinite number of solutions, because the rank of the matrix is less than its size (or determinant is 0). Each solution of such type of matrices is the product of a constant by an eigenvector of the matrix. This system can be solved in an analytic way by substituting the expression for one of the unknown amplitudes with another. Parameter β^2 is hidden in the

expressions for $k^x_{1.i}$ and $k^x_{2.i}$. The equation system written for the TE case can be reduced to

$$\cos(k_1 d_1)\cos(k_2 d_2) - \frac{1}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right)\sin(k_1 d_1)\sin(k_2 d_2) = \cos(^i k^x d) \qquad (3.4)$$

Writing down conditions for the TM mode we will get

$$\begin{cases} f_1 \exp(jk^x_{1.i}d_1) + f_2 \exp(-jk^x_{1.i}d_1) = f_3 \exp(jk^x_{2.i}d_1) + f_4 \exp(-jk^x_{2.i}d_1) \\ \frac{1}{\varepsilon_1}\left(k^x_{1.i} f_1 \exp(jk^x_{1.i}d_1) - k^x_{1.i} f_2 \exp(-jk^x_{1.i}d_1)\right) = \frac{1}{\varepsilon_2}\left(k^x_{2.i} f_3 \exp(jk^x_{2.i}d_1) - k^x_{2.i} f_4 \exp(-jk^x_{2.i}d_1)\right) \\ f_3 \exp(jk^x_{2.i}d) + f_4 \exp(-jk^x_{2.i}d) = f_1 \exp(jk^x_{1.i}d) + f_2 \exp(-jk^x_{1.i}d) \\ \frac{1}{\varepsilon_1}\left(k^x_{2.i} f_3 \exp(jk^x_{2.i}d) - k^x_{2.i} f_4 \exp(-jk^x_{2.i}d)\right) = \frac{1}{\varepsilon_2}\left(k^x_{1.i} f_1 \exp(jk^x_{1.i}d) - k^x_{1.i} f_2 \exp(-jk^x_{1.i}d)\right) \\ f_1 \exp(jk^x_{1.i}d) + f_2 \exp(-jk^x_{1.i}d) = \left[f_1 \exp(jk^x_{1.i}0) + f_2 \exp(-jk^x_{1.i}0)\right] \exp(j^i k^x d) \end{cases}$$

leads to the dispersion equation of the form:

$$\cos(k_1 d_1)\cos(k_2 d_2) - \frac{1}{2}\left(\frac{\varepsilon^g_2 k_1}{\varepsilon^g_1 k_2} + \frac{\varepsilon^g_1 k_2}{\varepsilon^g_2 k_1}\right)\sin(k_1 d_1)\sin(k_2 d_2) = \cos(^i k^x d) \quad (3.5)$$

On the right side of both equations (3.4, 3.5) is a value defined by the incident wave projection and the grating period, the left hand side includes only "internal" grating parameters. The left-hand side can be considered as the dispersion function, characterised only by the grating. Graphically these functions in the β^2 axis will look like depicted in fig.3.3 for TM case: The right side of the equation does not exceed 1 by modulus (assuming that the grating is irradiated with a propagating wave). Thus, the solution of the dispersion equation consists in finding the intersection points of two curves: a constant value corresponding to the right side (irradiation condition) and a curve corresponding to the left part of the eq.s (3.4), (3.5) (grating properties).

In the case of dielectric lossless gratings, all solutions belong to real β^2 numbers (lie on the real axis). This makes it easier to find the distribution of the propagating constants. In the case of metals, as was shown in [46] the problem is not so simple and there are "hidden" modes: solutions that satisfy the dispersion equation, but lying in the complex domain of β^2. All hidden modes are coupled: they differ in the sign of the complex part of the β^2 value.

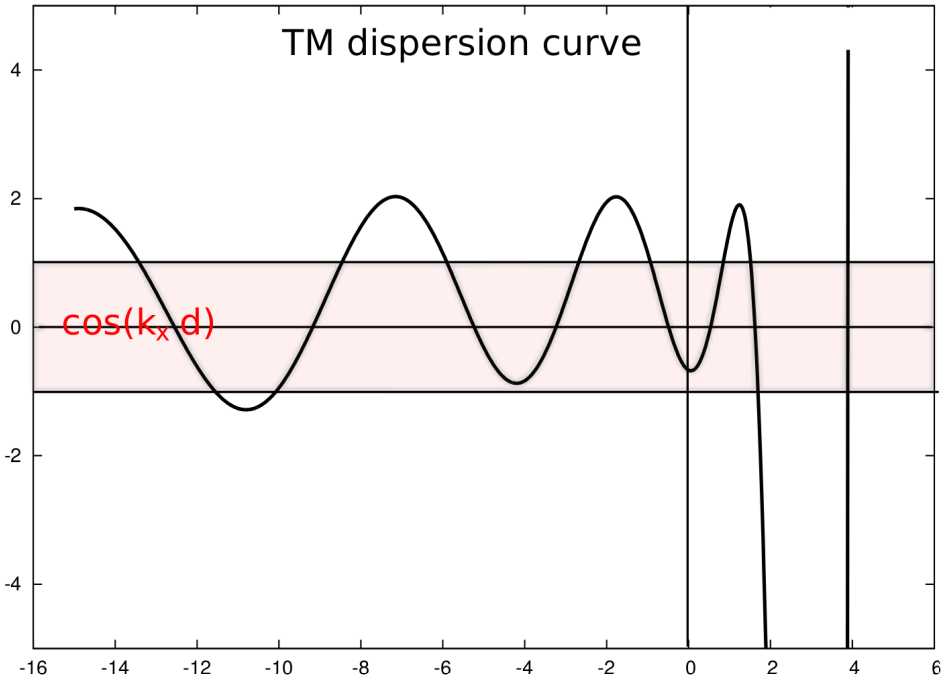


Figure 3.3: Dispersion curve example for the TM case as a function of β^2

3.6 Generalisation to the case of gratings composed of more than two sections

The task of finding the propagation constants can be generalised to a larger number of sections over the period. For any number n (number of the components over the period) we can write $2 \cdot n$ equations to the field amplitudes and their derivatives on each pair of boundaries. So there will be $2 \cdot n$ equations to determine the amplitudes and $2 \cdot n$ unknown amplitudes and still

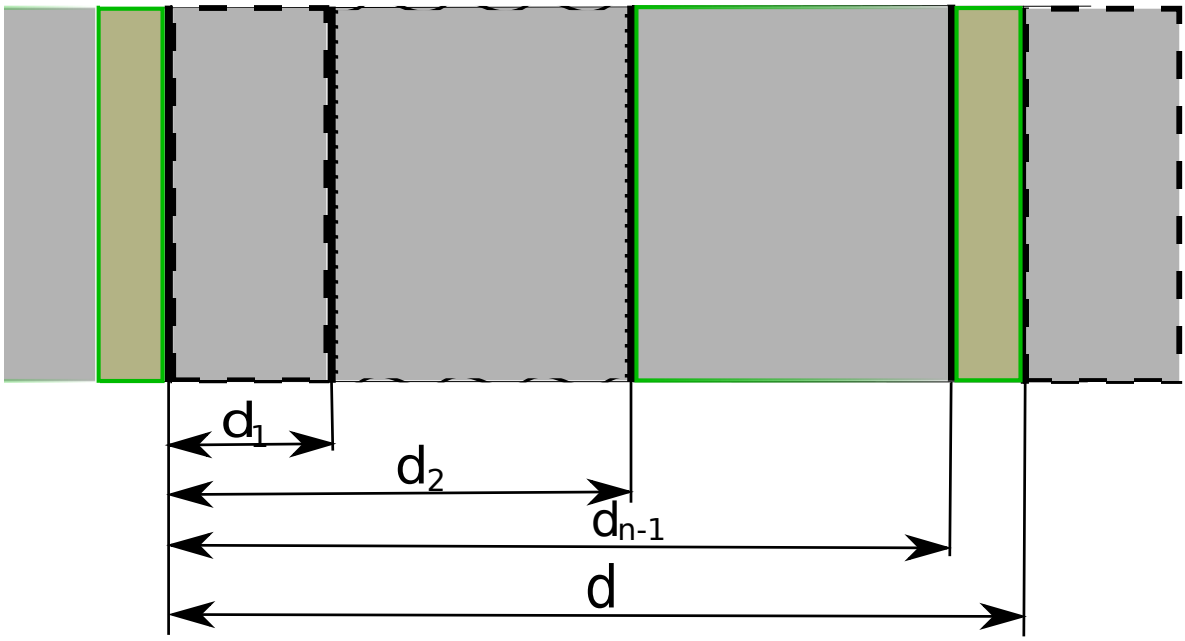


Figure 3.4: Notation for grating composed of more then two materials

there would be one more unknowns than the equations (β^2).

$$\begin{cases} \psi_1(d_1-0) & = & \psi_2(d_1+0) \\ \psi_1'(d_1-0) & = & \psi_2'(d_1+0) \\ \psi_2(d_2-0) & = & \psi_3(d_2+0) \\ \psi_2'(d_2-0) & = & \psi_3'(d_2+0) \\ \cdots & & \\ \psi_{n-1}(d_{n-1}-0) & = & \psi_n(d_{n-1}+0) \\ \psi_{n-1}'(d_{n-1}-0) & = & \psi_n'(d_{n-1}+0) \\ \psi_n(d-0) & = & \psi_1(d+0) \\ \psi_n'(d-0) & = & \psi_1'(d+0) \\ \psi_i(x) & = & \psi_i(x+d)\exp(j\dot{k}_x d) \end{cases}$$

Introducing partially defined $\psi(x)$ function:

$$\psi(x) = \begin{cases} {}^1\psi(x) \\ \cdots \\ {}^n\psi(x) \end{cases} = \begin{cases} fp_1 \exp(j^x k^1 x) + fn_1 \exp(-j^x k^1 x) & ,0 \le x \le d_1 \\ fp_2 \exp(j^x k^2 x) + fn_2 \exp(-j^x k^2 x) & d_1 \le x \le d_2 \\ \cdots & \cdots \\ fp_n \exp(j^x k^n x) + fn_n \exp(-j^x k^n x) & d_{n-1} \le x \le d \end{cases}$$

The dispersion equation in this case can be rewritten in a matrix form:

$$\begin{pmatrix} \Theta_1^+ & \Theta_2^- & 0 & \cdots & 0 & 0 \\ 0 & \Theta_2^+ & \Theta_3^- & \cdots & 0 & 0 \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & \Theta_{n-1}^+ & \Theta_n^- \\ \Theta_1^- & 0 & 0 & \cdots & 0 & \Theta_n^+ \end{pmatrix}$$
$$\times (fp_1, fn_1, fp_2, fn_2, \cdots, fp_{n-1}, fn_{n-1}, fp_n, fn_n)^T = 0$$

where Θ_i^+ is the coefficient matrix written for the right border of the slab i:

$$\Theta_i^+ = \begin{pmatrix} \exp(jk_i d_i) & \exp(-jk_i d_i) \\ k_i \frac{1}{\delta_i} \exp(jk_i d_i) & -k_i \frac{1}{\delta_i} \exp(-jk_i d_i) \end{pmatrix}$$

Θ_i^- is the coefficient matrix written for the left border of the slab i:

$$\Theta_i^- = \begin{pmatrix} -\exp(jk_i d_{i-1}) & -\exp(-jk_i d_{i-1}) \\ -k_i \frac{1}{\delta_i} \exp(jk_i d_{i-1}) & k_i \frac{1}{\delta_i} \exp(-jk_i d_{i-1}) \end{pmatrix}$$

and

$$\Theta_1^- = \begin{pmatrix} -\exp(jk_1 0) & -\exp(-jk_1 0) \\ -k_1 \frac{1}{\delta_1} \exp(jk_1 0) & k_1 \frac{1}{\delta_1} \exp(-jk_1 0) \end{pmatrix} \times \exp(j^i k^x d)$$

δ_i is the permittivity of the i-th slab for TM modes, and permeability of the slab for TE case. Resolving that matrix equation as a function of β^2 values would give propagation constants values, when

$$\det M(\beta^2) = 0$$

condition is satisfied.

It does not have the same clarity and is not straight-forward to understand, as in the binary grating case, but that equation possesses all the same properties that allow to distinguish the grating properties and the condition imposed by the incident wave.

In all the cases, the entries of the β into the equation terms are in square. It is dictated by Mathematics thus far, that sign of the β does not influence the resulting equations and solutions. On the other hand, we can also find some physical interpretation. We construct a basis in which the solutions will be distributed in both directions relative to the axis z. The field distribution and the properties of the modes do not depend on the propagation direction.

3.7 Normalisation of modes

After deriving the propagation constants it is possible to build the distribution of the modal functions along the period. One can put one of the amplitudes equal to 1 and calculate the remaining components. That operation is equal to calculating of the matrix's eigenvector. Such approach would show a distribution of a mode along the period but it would not be possible to compare them between each other because each of the eigenvector is defined up to a constant. For comparison purposes it is necessary to normalise modes according to the same rule. It is easy to reach that goal by introducing an orthogonality operator or an orthogonality criteria. For a complex differential operator of the dispersion equation it is possible to construct a conjugate. Development of that idea can be found in [many-many links]. We can find in these papers that in the simple case of materials without losses, simple complex conjugation of the dispersion function would provide the orthogonal operator. For problems with lossy materials there should be a more strict approach used. The orthogonal operator satisfies the following condition:

$$< L(\psi), \overline{L(\overline{\psi})} >= 0$$

where overline means conjugation, and brackets $< f(x) >= \frac{1}{d}\int_0^d f(x)dx$ stands for averaging over one period. From the orthogonality of the operators orthogonality of the functions follow.

Thus, the conjugated operator is orthogonal to the operator of the original problem and there is a correspondence for each function of original to the orthogonal function

$$\int_0^d \frac{\psi_i(x)\left(\overline{\psi_j(x)}\right)^*}{\omega\delta(x)}dx = \begin{cases} 0, i \neq j \\ v, i = j \end{cases}$$

where $\delta(x)$ means permittivity for the TM modes and permeability in TE case, and v is some arbitrary value. This feature can be used for the normalisation of the modes: calculating the integral over one period and setting it equal to one:

$$\int_0^d \frac{\psi_i(x)\overline{\psi_i(x)}^*}{\omega\delta(x)}dx = 1$$

Function $\psi(x)$ is partially defined and amplitudes in each section are defined up to some constant. Setting arbitrary amplitudes of the partially defined function ψ we would obtain a finite number C^2, not equal to zero:

$$\int_0^d \frac{\psi_i(x)\overline{\psi_i(x)}^*}{\omega\delta(x)}dx = C^2$$

from where we can find normalised modal function $\psi_n(x)$ if we got the modal function $\psi(x)$ and corresponding to that the function value C:

$$\psi_n(x) = \frac{\psi(x)}{C}$$

There remains an ambiguity in the phase of each mode with respect to each other for a certain problem. Way of solving diffraction problem by combining scattering matrix from two transition matrix avoids explicit calculation of the modal amplitudes but after the problem is solved it is possible

to calculate amplitudes and phases of each mode.

Here are examples of the several modes, distributed along the period of the structure. The white region is with refraction index $n = 1$, coloured region with $n = 2.2$ for the dielectric case and $\varepsilon = -4.84$ for the metal case. The distributions are obtained for normal incidence of the light with wavelength $\lambda = 628nm$, grating period $\Lambda = 400nm$ and filling factor $mr = 0.48$.

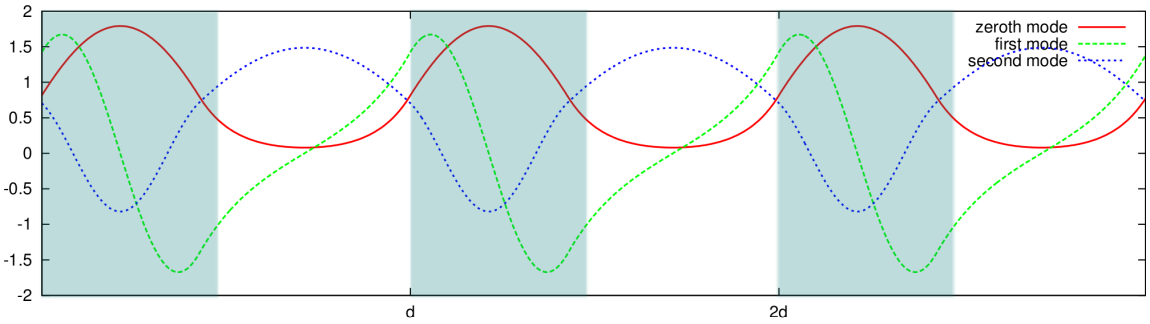


Figure 3.5: First three TE mode distributions along the grating in dielectric case

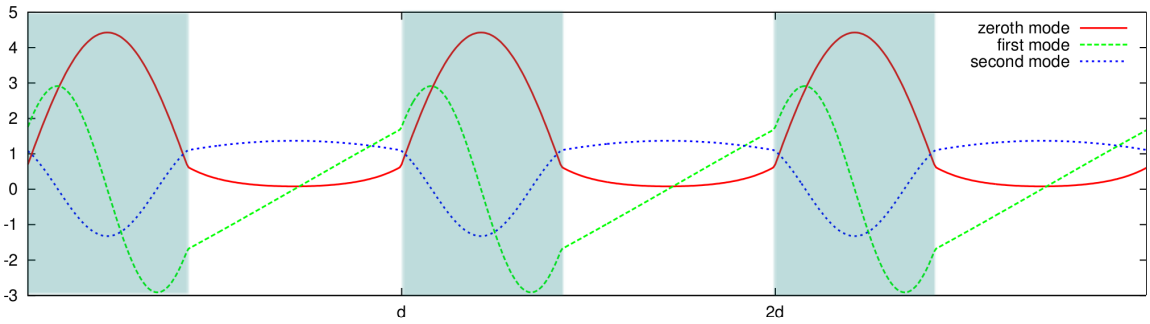


Figure 3.6: First three TM mode distributions along the grating in dielectric case

3.8 Mode-order compliment

As the number of the mode increases we can notice that the modal field distribution becomes more and more like the diffraction order's one. Amplitudes of each mode decrease. This similarity will be mathematically explained in

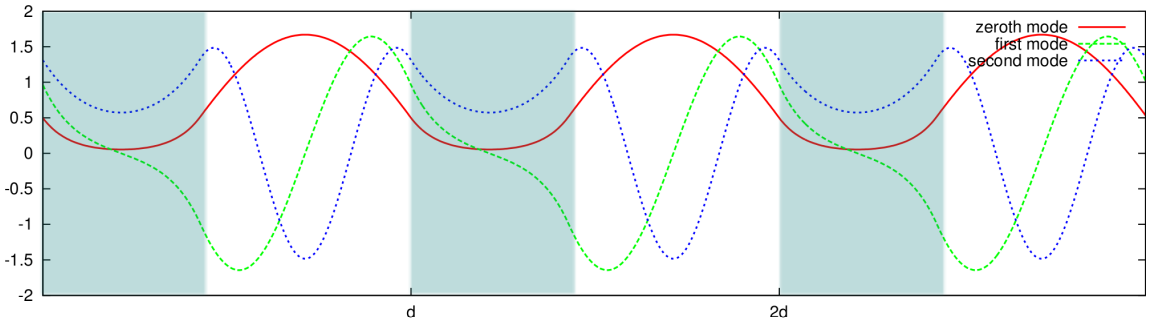


Figure 3.7: First three TE mode distributions along the grating in metal case

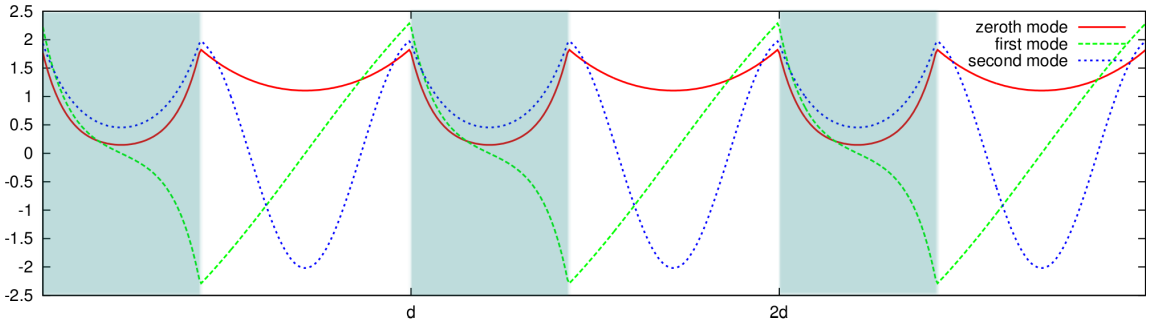


Figure 3.8: First three TM mode distributions along the grating in metal case

the next section. We may conclude from the pictures, that there is a correspondence of the orders and modes. We may expect that energy transmitted with a certain order would be more concentrated in the corresponding mode and vice-verse. For low numbered modes correspondence with the orders can be not so strong and evident, but for some particular cases such correspondence can be observed. This will be also expressed in the following section based on the overlap integrals.

Metallic grating does not completely follow such behaviour. However the correspondence between orders and modes also can be found.

3.9 The overlap integrals

At this stage of a model construction it is possible to express any field solution within the grating by a superposition of modes where each of them is

the Maxwell's equations solution and orthogonal (in general sense) to others. However, the amplitude of each mode, that would be excited by the incident beam, is still uncertain. It is also clear, that the transmitted and reflected energy of the propagating electromagnetic wave must be conserved for structures without losses. Based on this simple understanding of the field propagation, we would express the field at each of the boundaries of the grating (boundary where the light comes from and the opposite one). Without limiting the generality of the problem, consider the illumination of the grating by the several beams, so that for k_x projection of the beam the following expression would be valid:

$$k_n^x = k_0^x + n\Lambda$$

Each wave can have its own amplitude and polarisation. In this case, the incident field is most conveniently described as a vector of the amplitudes of each of the orders and the corresponding polarisations. A full field at the boundary will be expressed as the superposition of incident waves:

$$\begin{aligned}
E_x &= \sum_{n=-\infty}^{\infty} \quad \tfrac{k^y}{\gamma_n} \cdot A_n^e \quad -\tfrac{k_n^z}{\omega\varepsilon(x)} \tfrac{k_n^x}{\gamma_n} \cdot A_n^h \\
E_y &= \sum_{n=-\infty}^{\infty} \quad -\tfrac{k_n^x}{\gamma_n} \cdot A_n^e \quad -\tfrac{k_n^z}{\omega\varepsilon(x)} \tfrac{k^y}{\gamma_n} \cdot A_n^h \\
E_z &= \sum_{n=-\infty}^{\infty} \quad 0 \qquad\quad +\tfrac{\gamma_n}{\omega\varepsilon(x)} \cdot A_n^h \\
H_x &= \sum_{n=-\infty}^{\infty} \quad \tfrac{k_n^z}{\omega\mu} \tfrac{k_n^x}{\gamma_n} \cdot A_n^e \quad +\tfrac{k^y}{\gamma_n} \cdot A_n^h \\
H_y &= \sum_{n=-\infty}^{\infty} \quad \tfrac{k_n^z}{\omega\mu} \tfrac{k^y}{\gamma_n} \cdot A_n^e \quad -\tfrac{k_n^x}{\gamma_n} \cdot A_n^h \\
H_z &= \sum_{n=-\infty}^{\infty} \quad \tfrac{\gamma_n}{\omega\mu} \cdot A_n^e \qquad\quad +0
\end{aligned} \tag{3.6}$$

where in notation A_n^p, p stands for the polarisation of the amplitude, n underscript stands for the order number and γ_n is the $n-th$ order projection on the grating surface plane. Full field outside the grating should be equal to the field inside the grating. Complete expression for the fields inside the grating are (3.3). When we equate order expressions with modal field rep-

resentation, truncating and taking only first orders in consideration ($2N+1$ orders) we would get following system (written only for x and y projections of the field) :

$$\begin{aligned}
\sum_{n=-N}^{N}\left(\frac{k^y}{\gamma_n}\cdot A_n^e-\frac{k_n^z}{\omega\varepsilon(x)}\frac{k_n^x}{\gamma_n}\cdot A_n^h\right)\exp(jk_n^x x) &= \sum_{i=0}^{M}\frac{\beta_i^h}{\omega\varepsilon(x)}a_i^h\psi_i^h(x)\\
\sum_{n=-N}^{N}\left(-\frac{k_n^x}{\gamma_n}\cdot A_n^e-\frac{k_n^z}{\omega\varepsilon(x)}\frac{k^y}{\gamma_n}\cdot A_n^h\right)\exp(jk_n^x x) &= \sum_{i=0}^{M}\left(\frac{k_i^{ze}}{\beta_i^e}a_i^e\psi_i^e(x)+\frac{1}{j\omega\varepsilon(x)}\frac{k^y}{\beta_i^h}a_i^h\psi_i^{h\prime}(x)\right)\\
\sum_{n=-N}^{N}\left(\frac{k_n^z}{\omega\mu}\frac{k_n^x}{\gamma_n}\cdot A_n^e+\frac{k^y}{\gamma_n}\cdot A_n^h\right)\exp(jk_n^x x) &= \sum_{i=0}^{M}-\frac{\beta_i^e}{\omega\mu}a_i^e\psi_i^e(x)\\
\sum_{n=-N}^{N}\left(\frac{k_n^z}{\omega\mu}\frac{k^y}{\gamma_n}\cdot A_n^e-\frac{k_n^x}{\gamma_n}\cdot A_n^h\right)\exp(jk_n^x x) &= \sum_{i=0}^{M}\left(-\frac{1}{j\omega\mu}\frac{k^y}{\beta_i^e}a_i^e\psi_i^{e\prime}(x)+\frac{k_i^{zh}}{\beta_i^h}a_i^h\psi_i^h(x)\right)
\end{aligned} \tag{3.7}$$

This equation relates the amplitudes from one side of the grating (orders) with the same number of modal amplitudes inside the grating. There are 2 options to express one set of amplitudes by another, to obtain an explicit dependence of each of the amplitudes from one set of amplitudes by the other side amplitudes. To obtain such an expression, we do the following: each of the equations is multiplied successively by the conjugate to the chosen distribution function and integrated over one period. We can multiply by conjugate function of the Rayleigh's orders. Rayleigh's orders basis has an orthogonal basis, which allows to perform such an operation with respect to any field.

$$\mathcal{O}(f(x))=\frac{1}{d}\int_0^d f(x)\cdot\exp(-jk_n^x x)dx$$

As a result we obtain $\times(2O+1)$ groups of 4 equations. Each group has only a specific order number on the left side, while the right side always contains some of the integrals of the following structure:

$$\begin{aligned}
{}^eA_m^n &= \int_0^d \exp(-jk_n^x x)\psi_m^e(x)dx & {}^hA_m^n &= \int_0^d \exp(-jk_n^x x)\psi_m^h(x)dx\\
{}^eB_m^n &= \int_0^d \exp(-jk_n^x x)\tfrac{\psi_m^e(x)'}{\omega\mu}dx & {}^hB_m^n &= \int_0^d \exp(-jk_n^x x)\tfrac{\psi_m^h(x)'}{\omega\varepsilon(x)}dx\\
{}^eC_m^n &= \int_0^d \exp(-jk_n^x x)\tfrac{\psi_m^e(x)}{\omega\mu}dx = \tfrac{{}^eA_n^m}{\omega\mu} & {}^hC_m^n &= \int_0^d \exp(-jk_n^x x)\tfrac{\psi_m^h(x)}{\omega\varepsilon(x)}dx
\end{aligned}$$

These expression are overlap integrals. Their values express how mode corresponds to the order, how they match each other. All the energy concentrated in the order will be distributed into the modes. Values correspond to the energy part of the order transmitted into the mode.

Using that set of conjugate functions allows to come to express the orders via modes. We note vector of modal amplitudes as $\vec{M}$ and vector of the order amplitudes as $\vec{O}$:

$$\vec{O} = T \times \vec{M}$$

T is transition matrix from the modes to the diffraction orders.

It is also possible to use set of functions orthogonal to the modal functions. This way by successive multiplication with orthogonal functions and integration over the period

$$\mathcal{P}(\circ) = \frac{1}{d} \int_0^d \circ \cdot \left(\overline{\psi_m^p(x)}\right)^* dx$$

we will come to the expression where left side is the sum of overlapping integrals, while on the right side there will be a sum of several modal amplitudes. Expressions for the overlap integrals are alike:

$$\begin{aligned}
{}^eA_n^m &= \int_0^d \left(\overline{\psi_m^e(x)}\right)^* \exp(jk_n^x x)dx & {}^hA_n^m &= \int_0^d \left(\overline{\psi_m^h(x)}\right)^* \exp(jk_n^x x)dx \\
{}^eB_n^m &= \int_0^d \frac{\left(\overline{\psi_m^{e\,\prime}(x)}\right)^*}{\omega\mu} \exp(jk_n^x x)dx & {}^hB_n^m &= \int_0^d \frac{\left(\overline{\psi_m^{h\,\prime}(x)}\right)^*}{\omega\varepsilon(x)} \exp(jk_n^x x)dx \\
{}^eC_n^m &= \int_0^d \frac{\left(\overline{\psi_m^e(x)}\right)^*}{\omega\mu} dx = \frac{{}^eA_n^m}{\omega\mu} & {}^hC_n^m &= \int_0^d \frac{\left(\overline{\psi_m^h(x)}\right)^*}{\omega\varepsilon(x)} \exp(jk_n^x x)dx
\end{aligned} \tag{3.8}$$

Using these expression would allow to write the matrix equation in the alternative form:

$$\vec{M} = \overline{T} \times \vec{O}$$

$\overline{T}$ matrix gives correspondence of the orders amplitude to the modal amplitudes.

3.10 The conical mount configuration

The resolution of the dispersion equation gives the propagation constants of modes satisfying the periodicity condition imposed by the incident wave projection perpendicular to the grating lines. The modal basis is independent of the wave vector projection along the grating line.

In case of the geometry provided in the beginning of this chapter, k_y does not influence the modal basis. For any value of k_y the constructed basis would satisfy the equation system. The modes field distribution would remain the same for any projection; propagation constants are also the same, but the distribution between z and y differ (y-projection is always equal to k_y). For projections will be valid:

$$\beta_i^2 = k^{y2} + k_i^{z2}$$

where k_i^z depends on the β_i and k^y is the same for all the modes.

Additionally useful for building future bases in two-dimensional gratings is the fact that from a mathematical point of view there never was a constraint to the projection along the grating lines k^y. From a physical and practical point of view,it is impossible to radiate the grating so that the component of the incident wave is evanescent. But strict treatment of the problem does not impose any restriction on the projections of the wave vector. There is only a limitation to the vector sum of the components whereas the distribution of parts between components can be arbitrary.

This fact allows to use true-modal method as a basis for the two-dimensional true-modal method.

3.11 Transition matrix

Expressions obtained in the section 3.9 contain mathematical sums of the amplitudes. We can re-express them, so that the left side would contain only one amplitude. In case of expressing order amplitudes, that is quite trivial,

because

$$\frac{1}{d}\int_0^d \exp(jk_i^x x)\exp(-jk_j^x x)dx = \delta_{ij}$$

Expanding the equation system by separating waves going upwards and downwards we may write for a certain order number o (3.7):

$$\begin{array}{ll}
\frac{k^y}{\gamma_n} \times M_n^{+e} & -\frac{k_n^z}{\omega\varepsilon_b}\frac{k_n^x}{\gamma_n} \times M_n^{+h} + \frac{k^y}{\gamma_n} \times M_n^{-e} + \frac{k_n^z}{\omega\varepsilon_b}\frac{k_n^x}{\gamma_n} \times M_n^{-h} = \\
 & \sum_{i=0}^{M} \left(\beta_i^h a_i^{+h} + \beta_i^h a_i^{-h}\right) {}^hC_i^o \\
-\frac{k_n^x}{\gamma_n} \times M_n^{+e} & -\frac{k_n^z}{\omega\varepsilon_b}\frac{k^y}{\gamma_n} \times M_n^{+h} - \frac{k_n^x}{\gamma_n} \times M_n^{-e} + \frac{k_n^z}{\omega\varepsilon_b}\frac{k^y}{\gamma_n} \times M_n^{-h} = \\
 & \sum_{i=0}^{M} \frac{k_i^{ze}}{\beta_i^e}\left(a_i^{+e} - a_i^{-e}\right) {}^eA_i^o - j\frac{k^y}{\beta_i^h}\left(a_i^{+h} + a_i^{-h}\right) {}^hB_i^o \\
\frac{k_n^z}{\omega\mu}\frac{k_n^x}{\gamma_n} M_n^{+e} & +\frac{k^y}{\gamma_n} \times M_n^{+h} - \frac{k_n^z}{\omega\mu}\frac{k_n^x}{\gamma_n} M_n^{-e} + \frac{k^y}{\gamma_n} \times M_n^{-h} = \\
 & \sum_{i=0}^{M} -\frac{\beta_i^e}{\omega\mu}\left(a_i^{+e} + a_i^{-e}\right) {}^eA_i^o \\
\frac{k_n^z}{\omega\mu}\frac{k^y}{\gamma_n} \times M_n^{+e} & -\frac{k_n^x}{\gamma_n} \times M_n^{+h} - \frac{k_n^z}{\omega\mu}\frac{k^y}{\gamma_n} \times M_n^{-e} - \frac{k_n^x}{\gamma_n} \times M_n^{-h} = \\
 & \sum_{i=0}^{M} \frac{k^y}{\beta_i^e}\left(a_i^{+e} + a_i^{-e}\right) {}^eB_i^o + \frac{k_i^{zh}}{\beta_i^h}\left(a_i^{+h} - a_i^{-h}\right) {}^hA_i^o
\end{array} \tag{3.9}$$

Here ε_b denotes the permittivity of the semi-infinite media at the boundary for which the equation is written. After some arithmetic (supposing that $\gamma_n \neq 0$ for each order number) we would arrive at the equation for each order number n:

$$\begin{aligned}
M_n^{+e} = \frac{1}{2\gamma_n}\sum_{i=0}^{M} & \\
& \left(-k_n^x \frac{k_i^{ze}}{\beta_i^e} {}^eA_i^n + \frac{\omega\mu}{k_n^z}\left(-k_n^x \frac{\beta_i^e}{\omega\mu} {}^eA_i^n + k^y \frac{k^y}{\beta_i^e} {}^eB_i^n\right)\right) a_i^{+e} \\
& + \left(k_n^x \frac{k_i^{ze}}{\beta_i^e} {}^eA_i^n + \frac{\omega\mu}{k_n^z}\left(-k_n^x \frac{\beta_i^e}{\omega\mu} {}^eA_i^n + k^y \frac{k^y}{\beta_i^e} {}^eB_i^n\right)\right) a_i^{-e} \\
& + k^y \left(\beta_i^h {}^hC_i^n + jk_n^x \frac{1}{\beta_i^h} {}^hB_i^n + \frac{\omega\mu}{k_n^z}\frac{k_i^{zh}}{\beta_i^h} {}^hA_i^n\right) a_i^{+h} \\
& + k^y \left(\beta_i^h {}^hC_i^n + jk_n^x \frac{1}{\beta_i^h} {}^hB_i^n - \frac{\omega\mu}{k_n^z}\frac{k_i^{zh}}{\beta_i^h} {}^hA_i^n\right) a_i^{-h}
\end{aligned}$$

$$
\begin{aligned}
M_n^{+h} = \tfrac{1}{2\gamma_n} \sum_{i=0}^{M} & \\
& k^y \left(-\tfrac{\beta_i^e}{\omega\mu} {}^e A_i^n - \tfrac{k_i^{ze}}{\beta_i^e} \tfrac{\omega\varepsilon_b}{k_n^z} {}^e A_i^n - j \tfrac{k_n^x}{\beta_i^e} {}^e B_i^n \right) a_i^{+e} \\
& + k^y \left(-\tfrac{\beta_i^e}{\omega\mu} {}^e A_i^n + \tfrac{k_i^{ze}}{\beta_i^e} \tfrac{\omega\varepsilon_b}{k_n^z} {}^e A_i^n - j \tfrac{k_n^x}{\beta_i^e} {}^e B_i^n \right) a_i^{-e} \\
& - \left(\tfrac{\omega\varepsilon_b}{k_n^z} \left(\beta_i^h k_n^{xh} C_i^n - j k^y \tfrac{k^y}{\beta_i^h} {}^h B_i^n \right) + \tfrac{k_i^{zh}}{\beta_i^h} k_n^{xh} A_i^n \right) a_i^{+h} \\
& - \left(\tfrac{\omega\varepsilon_b}{k_n^z} \left(\beta_i^h k_n^{xh} C_i^n - j k^y \tfrac{k^y}{\beta_i^h} {}^h B_i^n \right) - \tfrac{k_i^{zh}}{\beta_i^h} k_n^{xh} A_i^n \right) a_i^{-h}
\end{aligned}
$$

$$
\begin{aligned}
M_n^{-e} = \tfrac{1}{2\gamma_n} \sum_{i=0}^{M} & \\
& \left(-k_n^x \tfrac{k_i^{ze}}{\beta_i^e} {}^e A_i^n - \tfrac{\omega\mu}{k_n^z} \left(-k_n^x \tfrac{\beta_i^e}{\omega\mu} {}^e A_i^n + k^y \tfrac{k^y}{\beta_i^e} {}^e B_i^n \right) \right) a_i^{+e} \\
& + \left(k_n^x \tfrac{k_i^{ze}}{\beta_i^e} {}^e A_i^n - \tfrac{\omega\mu}{k_n^z} \left(-k_n^x \tfrac{\beta_i^e}{\omega\mu} {}^e A_i^n + k^y \tfrac{k^y}{\beta_i^e} {}^e B_i^n \right) \right) a_i^{-e} \\
& + k^y \left(\beta_i^{hh} C_i^n + j k_n^x \tfrac{1}{\beta_i^h} {}^h B_i^n - \tfrac{\omega\mu}{k_n^z} \tfrac{k_i^{zh}}{\beta_i^h} {}^h A_i^n \right) a_i^{+h} \\
& + k^y \left(\beta_i^{hh} C_i^n + j k_n^x \tfrac{1}{\beta_i^h} {}^h B_i^n + \tfrac{\omega\mu}{k_n^z} \tfrac{k_i^{zh}}{\beta_i^h} {}^h A_i^n \right) a_i^{-h}
\end{aligned}
$$

$$
\begin{aligned}
M_n^{-h} = \tfrac{1}{2\gamma_n} \sum_{i=0}^{M} & \\
& k^y \left(-\tfrac{\beta_i^e}{\omega\mu} {}^e A_i^n + \tfrac{k_i^{ze}}{\beta_i^e} \tfrac{\omega\varepsilon_b}{k_n^z} {}^e A_i^n - j \tfrac{k_n^x}{\beta_i^e} {}^e B_i^n \right) a_i^{+e} \\
& + k^y \left(-\tfrac{\beta_i^e}{\omega\mu} {}^e A_i^n - \tfrac{k_i^{ze}}{\beta_i^e} \tfrac{\omega\varepsilon_b}{k_n^z} {}^e A_i^n - j \tfrac{k_n^x}{\beta_i^e} {}^e B_i^n \right) a_i^{-e} \\
& + \left(-\tfrac{k_i^{zh}}{\beta_i^h} k_n^{xh} A_i^n + \tfrac{\omega\varepsilon_b}{k_n^z} \left(\beta_i^h k_n^{xh} C_i^n - j k^y \tfrac{k^y}{\beta_i^h} {}^h B_i^n \right) \right) a_i^{+h} \\
& + \left(\tfrac{k_i^{zh}}{\beta_i^h} k_n^{xh} A_i^n + \tfrac{\omega\varepsilon_b}{k_n^z} \left(\beta_i^h k_n^{xh} C_i^n - j k^y \tfrac{k^y}{\beta_i^h} {}^h B_i^n \right) \right) a_i^{-h}
\end{aligned}
$$

The equations can be grouped like:

$$
\begin{pmatrix} M^{+e} \\ M^{+h} \\ M^{-e} \\ M^{-h} \end{pmatrix} = \begin{bmatrix} C_1 & k^y C_2 & \overline{C_1} & k^y \overline{C_2} \\ k^y C_3 & C_4 & k^y \overline{C_3} & \overline{C_4} \\ D_1 & k^y D_2 & \overline{D_1} & k^y \overline{D_2} \\ k^y D_3 & D_4 & k^y \overline{D_3} & \overline{D_4} \end{bmatrix} \times \begin{pmatrix} a^{+e} \\ a^{+h} \\ a^{-e} \\ a^{-h} \end{pmatrix}
$$

where C_i, $\overline{C_i}$, D_i and $\overline{D_i}$ are some terms expressed earlier. Matrix written in this form shows how the polarisations are separated in case of non-conical mount, when $k^y = 0$. It is possible to split the matrix into two separated matrices, each of which corresponds to particular polarisation.

Similar expression can be written for another boundary. This gives two transition matrices, which express order amplitudes at both boundaries of

the grating by unknown amplitudes of the modes.

We can also make similar derivation to obtain expression of the modes with order amplitudes. For infinite number of orders and modes (without truncation) we can write:

$$T = \left(\overline{T}\right)^{-1}$$

where T and $\overline{T}$ are transition matrices introduced in the section 3.9. For the finite matrices each of them should be calculated separately and the expression above is not fulfilled.

3.12 The scattering matrix of one layer

Possessing two T matrices at each grating boundary allows to calculate the scattering matrix according to the procedure, described in sec.2.2.2.

There are different definition points used in the literature for the modes. Some paper define modal amplitudes at the $z = 0$ point, and place the grating, so that bottom and top are equally distanced on $\frac{h}{2}$ from the $z = 0$ plane. In another papers, modes are counted starting from the boarder: mods going upwards are defined at the bottom and downwards modes are defined at the top of the grating. Amplitudes of the modes should be multiplied with corresponding phases (depending on the depth and modal vector projection): $\Psi_m = \exp(\pm j k_m^z h)$ where h is distance from the definition point to the border on interest, and k_m^z is the modal propagation constant projection on the $z-$axis.

As a result, the final scattering matrix of the one-dimensional matrix is obtained.

3.13 The transition grating-grating matrix

Sometimes, it is necessary to calculate the scattering matrix of a complex structure, composed of several gratings (stack of gratings). Resulting scattering matrix can be obtained by multiplication of the scattering matrices of each slice. This approach implicitly uses an order representation of the field

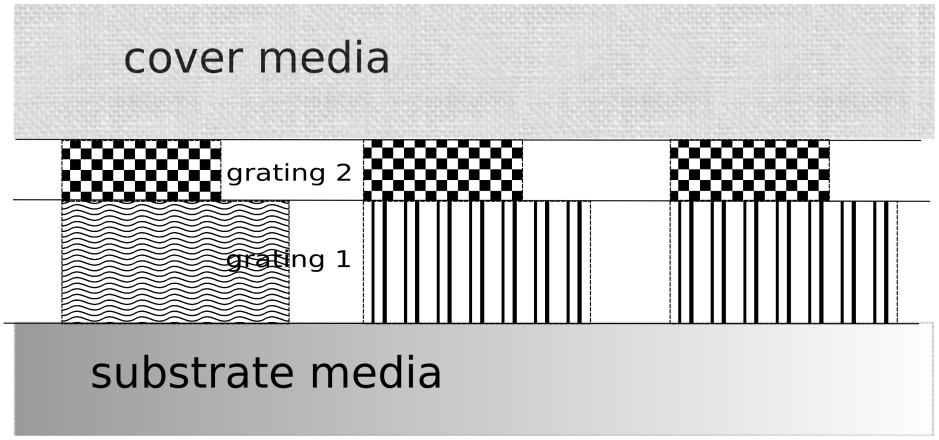


Figure 3.9: Complex structure example

between the slices in the semi-infinite media of zero-depth.

We may calculate the direct transition matrix for the grating-grating boundary. This approach decreases the number of calculation and brings benefits to the computation time needed for complex structures, calculated with slicing technique.

This technique uses the conjugate basis of the grating. Similar to the constructing of the $\overline{T}$ matrix, we apply orthogonal operator to both sides of the equation, putting the relation between modal fields inside one grating with the fields in another grating. Analytical expression of modal basis in both regions allows to derive overlap integrals analytically and reduce computation difficulty of the problem.

3.14 Conclusion

This chapter contains the summary on the true-modal method up to present dates. Analytical expressions allows to use this method as a basis for the two-dimensional method development. This method is well developed and there can be done just a few things concerning optimisation. There is still an open question for the fast propagation constants search. Each developer of the True-modal method sticks into this problem. It is possible either to find quickly part of the constants but lose accuracy or find all the propagation

constants with significant overhead during search. That problem mainly concern mathematical problem then physical part of the problem. True-modal 1D method can be used as a reference method for those who develops fast non-rigorous methods.

Chapter 4

Development of the Fourier modal method for one-dimensional grating with slanted walls

The proper field representation over a period is necessary for the constructing of the two-dimensional modal basis. In case when the grating can not be correctly represented as a stack of lamellar gratings, the technique, described in this chapter, can be applied to reduce computation time and remain accuracy of the results.

4.1 Introduction

The Fourier Modal Method (FMM)[47],[3] (also known as the Rigorous Coupled Wave Analysis (RCWA) as reactualised by Moharam and Gaylord [48] has been widely used in diffraction modelling due to simplicity of its implementation. This method is well developed for 1D binary grating for both TE and TM polarisations. Applying the FMM to gratings of arbitrary profile requires a representation of the actual profile in the form of a staircase approximation where the corrugation is represented as a stack of ultra-thin

binary gratings of different duty cycle with vertical walls [49]. It was shown recently [17] that such staircase approximation leads in the case of the TM polarisation to artifacts that do not represent the reality. Nevière and Popov compared the results obtained by their differential method modified to account for slanted boundaries at the edge of the slices [50] with those obtained by the FMM and demonstrated a definite accuracy improvement.

We found that similar modification can be applied to the FMM as well and that this gives rise to a notable improvement of its convergence. Section 2 gives the expression of the modified equations. The scattering S matrix is calculated for each slice with different oblique boundaries. The S matrix of full structure is then obtained by combining the matrices of all slices [51]. It is clear that using vertical boundaries instead of slanted when calculating S matrices is a very rough approximation and leads to calculation errors and poor convergence. Our technique represents all boundary conditions correspondingly to the angles given by the grating profile. This explains why calculation errors are smaller and the convergence is faster in our case.

It is worth to note here that there are several recent publications on the development of the FMM in the case of binary 2D gratings with perfectly normal walls [15, 52, 53]. But the authors of these articles do not consider relief gratings with oblique walls. Their analysis is limited to normal walls only whereas we consider the general case of arbitrarily profiled grating grooves.

We perform the comparison of different techniques in Section 3 to demonstrate the importance of correct representation of boundary conditions in the reciprocal space. The results are compared with those obtained by reference methods best suitable for the considered profiles. The precision is tested in a binary corrugation in comparison with the reference true-mode method and demonstrated in the case of a sinusoidal profile against the C-method [1].

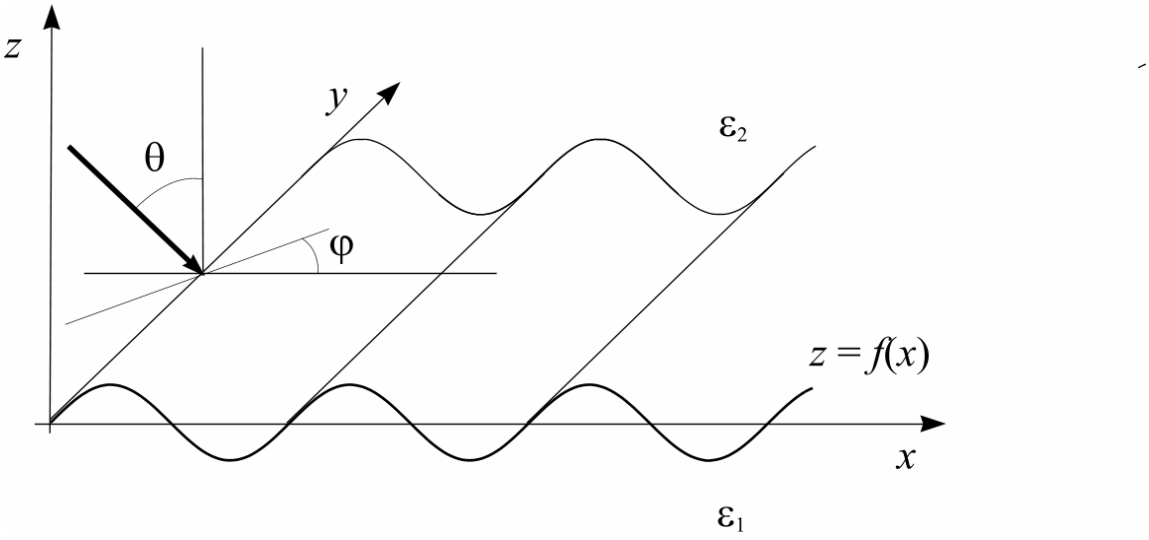


Figure 4.1: Problem geometry: conical incidence on a periodically modulated interface

4.2 FMM theoretical extension

4.2.1 Problem definition

Consider a 1D grating represented by a layer of permittivity $\varepsilon(x)$ periodically varying along the y axis with period d. In many practical cases function $\varepsilon(x)$ is discontinuous, i.e., it may comprise interfaces between different media composing the grating. Fig.4.1 represents the simple case of a periodically modulated interface in a region located between two semi-infinite media of absolute permittivity ε_1 and ε_2 .

The surface modulation is described by periodic function $f(x) = f(x+d)$. The grating plane is perpendicular to the z axis. The grating lines are aligned along the y-direction. A monochromatic plane wave is incident from the semi-infinite medium 1 under angles Θ and ϕ. Its electric and magnetic fields are proportional to factor $\exp(j\vec{\mathbf{k}}_0 \cdot \vec{\mathbf{r}})$ with

$$\begin{aligned} k_{0x} &= k_1 \sin(\Theta)\cos(\phi) \\ k_{0y} &= k_1 \sin(\Theta)\sin(\phi) \\ k_{0z} &= k_1 \cos(\Theta) \end{aligned} \tag{4.1}$$

$$k_i = \omega\sqrt{\mu_0 \varepsilon_i} \tag{4.2}$$

The temporal dependence factor $\exp(-j\omega t)$ is the same in all expressions; it will be omitted hereunder for brevity.

4.2.2 Modal development

We look for solutions of Maxwell's equations in the grating region

$$\begin{aligned} \nabla \times \vec{\mathbf{E}} &= j\omega\mu_0 \vec{\mathbf{H}} \\ \nabla \times \vec{\mathbf{H}} &= -j\omega \vec{\mathbf{D}} = -j\omega\varepsilon \vec{\mathbf{E}} \end{aligned} \tag{4.3}$$

applying the periodicity conditions

$$\begin{aligned} \vec{\mathbf{E}}(x+d, y, z) &= \vec{\mathbf{E}}(x, y, z)\exp(jk_{0x}d) \\ \vec{\mathbf{H}}(x+d, y, z) &= \vec{\mathbf{H}}(x, y, z)\exp(jk_{0x}d) \end{aligned} \tag{4.4}$$

The FMM defines its modal basis in the reciprocal Fourier space. Any solution to the diffraction problem in the grating region is represented by a superposition of modes propagating up and down along the z axis. If β is the modal propagation constant then, in accordance with (4.4), the modal field writes

$$\begin{aligned} \vec{\mathbf{E}}(x, y, z) &= \sum_{m=-\infty}^{\infty} \vec{\mathbf{E}}_m \exp(jk_{mx}x + jk_{0y} + j\beta z) \\ \vec{\mathbf{H}}(x, y, z) &= \sum_{m=-\infty}^{\infty} \vec{\mathbf{H}}_m \exp(jk_{mx}x + jk_{0y} + j\beta z) \end{aligned} \tag{4.5}$$

where $\vec{\mathbf{E}}_m$ and $\vec{\mathbf{H}}_m$ are field harmonics (depending on β and unknown at this stage),

$$k_{mx} = k_{0x} + m\frac{2\pi}{d}$$

Such field development is valid only if the grating structure possesses a translation symmetry along the z axis. In other words, it is supposed that the grating grooves are vertical and parallel to the yOz plane. This is true in

the case of binary (lamellar) gratings. Consequently, such binary gratings represent the basic structure of the modal method.

4.2.3 Slicing technique

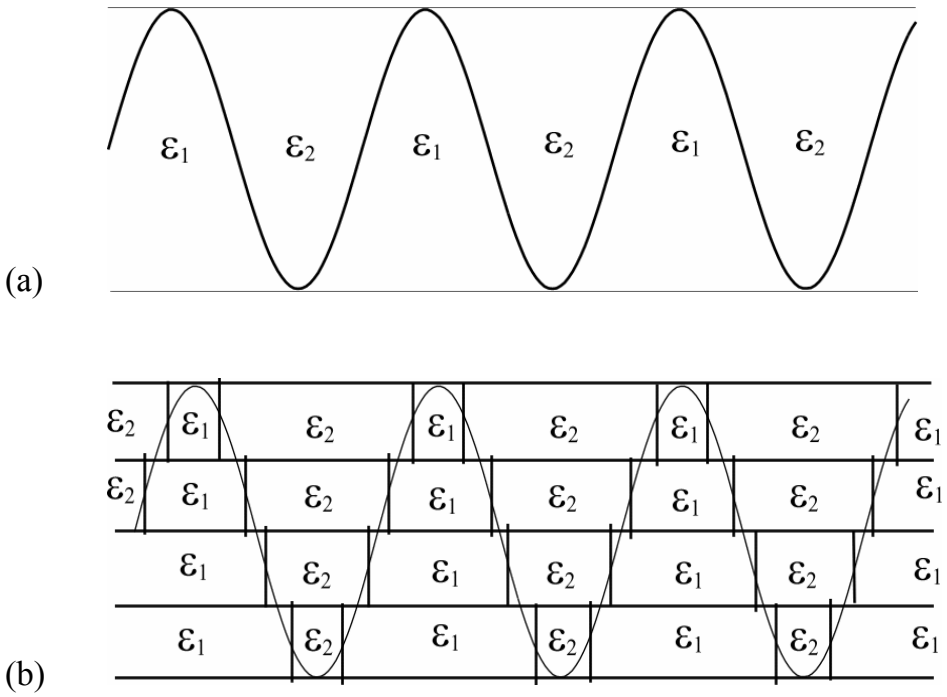


Figure 4.2: Approximation of the initial grating (a) by an array of lamellar gratings (b)

If the groove profile differs from the basic binary structure, the grating is approximated by a stack of thin lamellar gratings as sketched in fig.4.2. The diffraction S-matrix [51] is calculated for each elementary lamellar grating. Combining the S-matrices gives the S-matrix of the whole corrugation. This slicing technique rests on a stair-case approximation of the actual continuous profile; it is frequently used when the FMM is applied to gratings of arbitrary profile. At this stage one assumes a good convergence of solutions obtained to the exact solution in the limit of small slice thickness. The S-matrix recombination is a standard mathematical procedure and not much can be said at that stage. The most important issue here is calculation of the S

matrix for each individual grating slice. This is what the present section is devoted to.

Substituting solution (4.5) into (4.3) gives for each harmonics

$$\begin{aligned} k_{0y}E_{mz} - \beta E_{my} &= \omega\mu_0 H_{mx} \\ \beta E_{mx} - k_{mx}E_{mz} &= \omega\mu_0 H_{my} \\ k_{mx}E_{my} - k_{0y}E_{mx} &= \omega\mu_0 H_{mz} \end{aligned} \tag{4.6}$$

$$\begin{aligned} \beta H_{my} - k_{0y}H_{mz} &= \omega D_{mx} \\ k_{mx}H_{mz} - \beta H_{mx} &= \omega D_{my} \\ k_{0y}H_x - k_{mx}H_{my} &= \omega D_{mz} \end{aligned} \tag{4.7}$$

The transformation of the material relation $\vec{\mathbf{D}} = \varepsilon\vec{\mathbf{E}}$ into the reciprocal space is a key point in the FMM. In the reciprocal space the product of two periodical functions transforms to a convolution product:

$$\mathbf{D}_m = \sum_j [\varepsilon]_{mj} \mathbf{E}_j \tag{4.8}$$

where $[\varepsilon]$ is a Toeplitz matrix and its components $[\varepsilon]_{mj}$ are given by Fourier coefficients ε_{m-j} in the periodic permittivity representation by the Fourier series:

$$[\varepsilon]_{mj} = [\varepsilon]_{m,-j} = \frac{1}{d}\int_0^d \varepsilon(x) \exp\left[-i\frac{2\pi}{d}(m-j)x\right] dx$$

The summation in (4.8) is in principle infinite. Numerically however only a finite number of harmonics can be considered. Such truncation is the primary source of errors of the method. The crucial aspect here is the continuity of the functions involved. Permittivity $\varepsilon(x)$ is discontinuous at the boundaries. The Fourier transformation of the product of two distributions which are discontinuous at the very same point can not be performed according to rules. This means that expression (4.8) can only be applied for those components of the electric field which are continuous at the boundary. These are the

electric field components parallel to the boundary. It was shown in [42] how to deal with the normal component of the electric field $E_\perp$. The material relation has to be rewritten in the form $\frac{1}{\varepsilon}D_\perp = E_\perp$. The normal component of electric displacement $D_\perp$ is continuous at the boundary. Therefore, its product by the discontinuous inverse permittivity can be represented in the reciprocal space by the convolution:

$$\sum_j \left[\frac{1}{\varepsilon}\right]_{mj} D_{j\perp} = E_{m\perp}$$

where $\left[\frac{1}{\varepsilon}\right]$ is a Toeplitz matrix composed of corresponding coefficients:

$$\left[\frac{1}{\varepsilon}\right]_{mj} = [\varepsilon]_{m,-j} = \frac{1}{d}\int_0^d \frac{1}{\varepsilon(x)} \exp\left[-i\frac{2\pi}{d}(m-j)x\right]dx$$

Finally, the normal components of electric displacement are found by matrix inversion:

$$D_{m\perp} \sum_j \left[\frac{1}{\varepsilon}\right]^{-1}_{mj} E_{j\perp}$$

4.2.4 Slanted walls

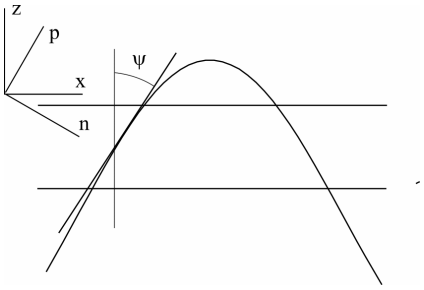


Figure 4.3: New coordinate system related to the slanted boundary

The stair-case approximation in its standard form [48], [49] is a quite rough approximation of the actual corrugation. The field calculated in each slice is likely to bear some characteristics of the lamellar geometry, especially in the case of the TM polarisation. A more exact field solution in each slice would be obtained if one could impose a new coordinate system on the field to satisfy the boundary conditions at the slanted boundary of the actual corrugation as suggested in fig.4.3.

To that end, we have developed a modified FMM formalism inspired from the general approach disclosed in [50] and applied to the differential method. While keeping the translation symmetry required by the FMM, a new orthogonal coordinate system attached to the actual corrugation profile is introduced as illustrated in fig. 4.3. It turns out that the FMM lends itself quite naturally to the integration of the oblique boundary conditions as will be seen hereunder. The y component of a vector is kept without changes but the two other components are turned:

$$\begin{aligned} E_n &= E_x \cos(\phi(x)) - E_z \sin(\phi(x)) \\ E_p &= E_x \sin(\phi(x)) + E_z \cos(\phi(x)) \end{aligned} \tag{4.9}$$

As $\phi(x)$ is the angle between the normal to the grating boundary in the XOZ plane (Fig.4.3), component E_n is normal and E_p is parallel to the boundary. The electric displacement field then is written in the new coordinate system:

$$\begin{aligned} D_{my} &= \sum_j [\varepsilon]_{mj} E_{jy} \\ D_{mp} &= \sum_j [\varepsilon]_{mj} E_{jp} \\ D_{mn} &= \sum_j \left[\frac{1}{\varepsilon}\right]^{-1}_{mj} E_{jn} \end{aligned} \tag{4.10}$$

In what follows, it is assumed that the grating's profile function $f(x)$ is differentiable. It follows that the continuous functions $\sin(\phi(x))$ and $\cos(\phi(x))$ can be defined from the derivative $f'(x)$:

$$\begin{aligned} \sin(\phi(x)) &= \frac{1}{\sqrt{1+[f'(x)]^2}} \\ \cos(\phi(x)) &= \frac{f'(x)}{\sqrt{1+[f'(x)]^2}} \end{aligned}$$

The products in relations (4.9) are transformed to convolution products:

$$\begin{aligned} E_{mn} &= \sum_j \left([c]_{mj} E_{jx} - [s]_{mj} E_{jz} \right) \\ E_{mp} &= \sum_j \left([s]_{mj} E_{jx} + [c]_{mj} E_{jz} \right) \end{aligned} \tag{4.11}$$

where $[s]$ and $[c]$ are Toeplitz matrices with coefficients

$$\begin{aligned} [s]_{mj} &= \frac{1}{d} \int_0^d \frac{1}{\sqrt{1 + [f'(x)]^2}} \exp\left[-i \frac{2\pi}{d} (m-j) x \right] dx \\ [c]_{mj} &= \frac{1}{d} \int_0^d \frac{f'(x)}{\sqrt{1 + [f'(x)]^2}} \exp\left[-i \frac{2\pi}{d} (m-j) x \right] dx \end{aligned}$$

The components of electric displacement are transformed similarly

$$\begin{aligned} D_{mx} &= \sum_j \left([c]_{mj} D_{jn} + [s]_{mj} D_{jp} \right) \\ D_{mx} &= \sum_j \left(-[s]_{mj} D_{jn} + [c]_{mj} D_{jp} \right) \end{aligned} \tag{4.12}$$

Substituting Eqs. (4.10), (4.11), and (4.12) into (4.6) gives new matrix relations between field components

$$\begin{aligned} \beta H_{my} - k_y H_{mz} &= \sum_j \omega \left([A]_{mj} E_{jx} + [B]_{mj} E_{jz} \right) \\ k_{mx} H_{mz} - \beta H_{mx} &= \sum_j \omega \left([\varepsilon]_{mj} E_{jy} \right) \\ k_y H_{mx} - k_{mx} H_{my} &= \sum_j \omega \left([C]_{mj} E_{jx} + [D]_{mj} E_{jz} \right) \end{aligned} \tag{4.13}$$

Matrices $[A]$,$[B]$,$[C]$ and $[D]$ from (4.13) are defined as

$$\begin{aligned} A &= [c]\left[\tfrac{1}{\varepsilon}\right]^{-1}[c] + [s][\varepsilon][s] \\ B &= [s][\varepsilon][c] - [c]\left[\tfrac{1}{\varepsilon}\right]^{-1}[s] \\ C &= [c][\varepsilon][s] - [s]\left[\tfrac{1}{\varepsilon}\right]^{-1}[c] \\ D &= [s]\left[\tfrac{1}{\varepsilon}\right]^{-1}[s] + [c][\varepsilon][c] \end{aligned}$$

Getting rid of z-projections of electric and magnetic fields yields the following matrix equation on unknown field amplitudes and modal propagation constant:

$$[M]\begin{bmatrix} E_x \\ E_y \\ H_x \\ H_y \end{bmatrix} = \beta \begin{bmatrix} E_x \\ E_y \\ H_x \\ H_y \end{bmatrix} \tag{4.14}$$

with

$$[M] = \begin{bmatrix} -[k_x][D]^{-1}[C] & 0 & \frac{k_y}{\omega}[k_x][D]^{-1} & \omega\mu_0[I] - \frac{1}{\omega}[k_x][D]^{-1}[k_x \\ -[k_y][D]^{-1}[C] & 0 & \frac{{k_y}^2}{\omega}[D]^{-1} - \omega\mu_0[I] & -\frac{k_y}{\omega}[D]^{-1}[k_x] \\ -\frac{k_y}{\omega\mu_0}[k_x] & \frac{1}{\omega\mu_0}[k_x]^2 - \omega[\varepsilon] & 0 & 0 \\ \omega[A] - \frac{{k_y}^2}{\omega\mu_0}[I] - \omega[B][D]^{-1}[C] & \frac{k_y}{\omega\mu_0}[k_x] & k_y[B][D]^{-1} & -[B][D]^{-1}[k_x] \end{bmatrix} \tag{4.15}$$

Here $[I]$ is a unit matrix and $[k_x]$ is the diagonal matrix with elements k_{mx}. Diagonalization of matrix M (4.15) gives modal propagation constants β_q together with the harmonic amplitudes of the modal fields. Note that in the present analysis up-propagating modes and down-propagating modes have, in general, different propagating constants. Moreover, in the conical geometry, both polarisations TE and TM are mixed in each mode. This doubles the final size of the matrix and, hence, the calculation time.

Under non-conical incidence,$k_y = 0$, and the matrix $[M]$ in Eq. 4.14 is split into two matrices giving independent solutions for TE modes:

$$\begin{bmatrix} 0 & -\omega\mu_0[I] \\ \frac{1}{\omega\mu_0}[k_x]^2 - \omega[\varepsilon] & 0 \end{bmatrix}\begin{bmatrix} E_y \\ H_x \end{bmatrix} = \beta \begin{bmatrix} E_y \\ H_x \end{bmatrix}$$

and for TM modes:

$$\begin{bmatrix} -[k_x][D]^{-1}[C] & \omega\mu_0[I] - \frac{1}{\omega}[k_x][D]^{-1}[k_x] \\ \omega[A] - \omega[B][D]^{-1}[C] & -[B][D]^{-1}[k_x] \end{bmatrix} \begin{bmatrix} E_y \\ H_x \end{bmatrix} = \beta \begin{bmatrix} E_y \\ H_x \end{bmatrix}$$

4.2.5 S matrix coefficients

The rest of the resolution scheme is a standard modal technique [23]. Although not much new is reported at this step, we have included it to complete the presentation of the novel FMM implementation. In the slice the electromagnetic solution is represented by a superposition of modal fields:

$$\begin{bmatrix} E_x \\ E_y \\ H_x \\ H_y \end{bmatrix} = \sum_q c_q \exp(j\beta_q z) \sum_m \begin{bmatrix} E_{qmx} \\ E_{qmy} \\ H_{qmx} \\ H_{qmy} \end{bmatrix} \exp(jk_{mx}x + jk_y y)$$

where c_q is the amplitude of the q-th order mode and $E_q m$ is the m-th harmonic amplitude of the electric field in its eigenvector. The homogeneous semi-infinite media below and above the grating have permittivities ε_b and ε_a, respectively. Without loss of generality, let us consider only the medium above the slice (the analysis at the lower interface is quite similar). The fields in the above medium are represented by a superposition of plane and evanescent TE and TM waves:

$$\begin{bmatrix} E_x \\ E_y \\ H_x \\ H_y \end{bmatrix} = \sum_m \exp(jk_{mx}x + jk_y y) \begin{bmatrix} E_{mx} \\ E_{my} \\ H_{mx} \\ H_{my} \end{bmatrix}$$

with components

$$\begin{bmatrix} E_{mx} \\ E_{my} \\ H_{mx} \\ H_{my} \end{bmatrix} = \begin{bmatrix} \frac{k_y}{\gamma_m} a^+_{me} - \frac{k_{mx}k^a_{mz}}{\omega\varepsilon_a\gamma_m} a^+_{mh} \\ \frac{k_{mx}}{\gamma_m} a^+_{me} - \frac{k_y k^a_{mz}}{\omega\varepsilon_a\gamma_m} a^+_{mh} \\ \frac{k_{mx}k^a_{mz}}{\omega\mu_0\gamma_m} a^+_{me} + \frac{k_y}{\gamma_m} a^+_{mh} \\ \frac{k_y k^a_{mz}}{\omega\mu_0\gamma_m} a^+_{me} - \frac{k_{mx}}{\gamma_m} a^+_{mh} \end{bmatrix} \exp(jk^a_{mz}z) + \begin{bmatrix} \frac{k_y}{\gamma_m} a^-_{me} + \frac{k_{mx}k^a_{mz}}{\omega\varepsilon_a\gamma_m} a^-_{mh} \\ -\frac{k_{mx}}{\gamma_m} a^-_{me} + \frac{k_y k^a_{mz}}{\omega\varepsilon_a\gamma_m} a^-_{mh} \\ -\frac{k_{mx}k^a_{mz}}{\omega\mu_0\gamma_m} a^-_{me} + \frac{k_y}{\gamma_m} a^-_{mh} \\ -\frac{k_y k^a_{mz}}{\omega\mu_0\gamma_m} a^-_{me} - \frac{k_{mx}}{\gamma_m} a^-_{mh} \end{bmatrix} \exp(-jk^a_{mz}z)$$

Coefficients a_{me}^{+} and a_{mh}^{+} are transverse field amplitudes of the TE and the TM diffraction waves, respectively (the "plus" sign corresponds to the upward propagating waves), and

$$\gamma_m = \sqrt{{k_{mx}}^2 + {k_y}^2}$$

$$k_{mz}^{a} = \sqrt{\omega^2 \mu_0 \varepsilon_a - \gamma_m^2}, \quad 0 \le \arg(k_{mz}^{a}) \le \pi$$

All the tangent field components are continuous at the top interface $z = f_{max}$. This delivers the following relations between modal amplitudes c_q and coefficients $a_m^{\pm}$:

$$\begin{bmatrix} a_{me}^{+} \\ a_{me}^{-} \\ a_{mh}^{+} \\ a_{mh}^{-} \end{bmatrix} = \sum_q \begin{bmatrix} \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & \frac{\omega\mu_0 k_{mx}}{2k_{mz}^{a}\gamma_m} & \frac{\omega\mu_0 k_y}{2k_{mz}^{a}\gamma_m} \\ \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & -\frac{\omega\mu_0 k_{mx}}{2k_{mz}^{a}\gamma_m} & -\frac{\omega\mu_0 k_y}{2k_{mz}^{a}\gamma_m} \\ -\frac{\omega\varepsilon_a k_{mx}}{2k_{mz}^{a}\gamma_m} & -\frac{\omega\varepsilon_a k_y}{2k_{mz}^{a}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \\ \frac{\omega\varepsilon_a k_{mx}}{2k_{mz}^{a}\gamma_m} & \frac{\omega\varepsilon_a k_y}{2k_{mz}^{a}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \end{bmatrix} \begin{bmatrix} E_{qmx} \\ E_{qmy} \\ H_{qmx} \\ H_{qmy} \end{bmatrix} c_q \exp\left[j(\beta_q \mp k_{mz}^{a})(z - f_{max})\right] \tag{4.16}$$

Similar consideration at the bottom interface $z = f_{min}$ gives:

$$\begin{bmatrix} b_{me}^{+} \\ b_{me}^{-} \\ b_{mh}^{+} \\ b_{mh}^{-} \end{bmatrix} = \sum_q \begin{bmatrix} \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & \frac{\omega\mu_0 k_{mx}}{2k_{mz}^{b}\gamma_m} & \frac{\omega\mu_0 k_y}{2k_{mz}^{b}\gamma_m} \\ \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & -\frac{\omega\mu_0 k_{mx}}{2k_{mz}^{b}\gamma_m} & -\frac{\omega\mu_0 k_y}{2k_{mz}^{b}\gamma_m} \\ -\frac{\omega\varepsilon_b k_{mx}}{2k_{mz}^{b}\gamma_m} & -\frac{\omega\varepsilon_b k_y}{2k_{mz}^{b}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \\ \frac{\omega\varepsilon_b k_{mx}}{2k_{mz}^{b}\gamma_m} & \frac{\omega\varepsilon_b k_y}{2k_{mz}^{b}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \end{bmatrix} \begin{bmatrix} E_{qmx} \\ E_{qmy} \\ H_{qmx} \\ H_{qmy} \end{bmatrix} c_q \exp\left[j(\beta_q \mp k_{mz}^{b})(z - f_{min})\right] \tag{4.17}$$

At the next step Eqs. (4.16) and (4.17) are resorted in two systems expressing the incoming and the outgoing wave amplitudes:

$$\begin{bmatrix} b^+_{me} \\ b^+_{mh} \\ a^-_{me} \\ a^-_{mh} \end{bmatrix} = \sum_q Q_{mq} c_q, \qquad \begin{bmatrix} b^-_{me} \\ b^-_{mh} \\ a^+_{me} \\ a^+_{mh} \end{bmatrix} = \sum_q R_{mq} c_q$$

where

$$Q_{mq} = \begin{bmatrix} \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & \frac{\omega\mu_0 k_{mx}}{2k^b_{mz}\gamma_m} & \frac{\omega\mu_0 k_y}{2k^b_{mz}\gamma_m} \\ -\frac{\omega\varepsilon_b k_{mx}}{2k^b_{mz}\gamma_m} & -\frac{\omega\varepsilon_b k_y}{2k^b_{mz}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \\ \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & -\frac{\omega\mu_0 k_{mx}}{2k^a_{mz}\gamma_m} & -\frac{\omega\mu_0 k_y}{2k^a_{mz}\gamma_m} \\ \frac{\omega\varepsilon_a k_{mx}}{2k^a_{mz}\gamma_m} & \frac{\omega\varepsilon_a k_y}{2k^a_{mz}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \end{bmatrix} \begin{bmatrix} E_{qmx} \exp\left[j(\beta_q - k^b_{mz})(z - f_{min})\right] \\ E_{qmy} \exp\left[j(\beta_q - k^b_{mz})(z - f_{min})\right] \\ H_{qmx} \exp\left[j(\beta_q + k^a_{mz})(z - f_{max})\right] \\ H_{qmy} \exp\left[j(\beta_q + k^a_{mz})(z - f_{max})\right] \end{bmatrix}$$

$$R_{mq} = \begin{bmatrix} \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & -\frac{\omega\mu_0 k_{mx}}{2k^b_{mz}\gamma_m} & -\frac{\omega\mu_0 k_y}{2k^b_{mz}\gamma_m} \\ \frac{\omega\varepsilon_b k_{mx}}{2k^b_{mz}\gamma_m} & \frac{\omega\varepsilon_b k_y}{2k^b_{mz}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \\ \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} & \frac{\omega\mu_0 k_{mx}}{2k^a_{mz}\gamma_m} & \frac{\omega\mu_0 k_y}{2k^a_{mz}\gamma_m} \\ -\frac{\omega\varepsilon_a k_{mx}}{2k^a_{mz}\gamma_m} & -\frac{\omega\varepsilon_a k_y}{2k^a_{mz}\gamma_m} & \frac{k_y}{2\gamma_m} & -\frac{k_{mx}}{2\gamma_m} \end{bmatrix} \begin{bmatrix} E_{qmx} \exp\left[j(\beta_q + k^b_{mz})(z - f_{min})\right] \\ E_{qmy} \exp\left[j(\beta_q + k^b_{mz})(z - f_{min})\right] \\ H_{qmx} \exp\left[j(\beta_q - k^a_{mz})(z - f_{max})\right] \\ H_{qmy} \exp\left[j(\beta_q - k^a_{mz})(z - f_{max})\right] \end{bmatrix}$$

The resulting S matrix of the grating slice is found as the inverse matrix product:

$$S = Q^{-1} R$$

4.3 Numerical examples of extended implementation of the FMM

The aim of this section is to demonstrate how different can be the results of implementation of the FMM depending on the technique used. The tests are performed in cases permitting an exact reference method. We analyzed convergence and final precision of the novel technique in comparison with previously reported versions of the FMM.

First we test results obtained for a binary grating with vertical walls by the true-mode method [46]. In this case $\phi = 0$, and we compare convergence of the Lalanne-Morris technique (LMT) [42] with that of the original FMM [3, 48]. Such comparison is not new. Similar results can be found elsewhere [42], [43], [54] showing the advantage of the LMT approach and the necessity of correct distributions multiplication. We perform this benchmark to precise typical rates of convergence for both techniques.

Next, we calculate a sinusoidal grating. Such gratings are well-known to be exactly solved by the C-method[1] which is used as a reference. At this step we show that simple application of the LMT technique with slicing leads to the same rate of convergence (or even worse) as the original FMM. We demonstrate that crucial improvement can be achieved for such profile applying the novel technique.

4.3.1 Dielectric lamellar grating

L.Li [43] analytically proved increasing of convergence rate for the LMT modification to the FMM [42]. But it looks impossible to find any detailed comparison of the convergence rates for these two techniques. Therefore, we perform such comparison of different techniques to fill this gap. We investigated convergence of the 0-th and the +1-st transmission order field amplitudes for the following structure: grating period d = 700 nm, filling factor 0.56, groove depth 460 nm, grating refractive index $n = 2.5$. The substrate is made of the same material as the grating. The second medium of the grating and the cover is air with $n = 1$. Normal incidence of a plane monochromatic wave $\lambda = 632.8$ nm with magnetic field aligned along grating lines (TM polarisation) is considered. Applying the true-mode method [46] and taking into account 1025 modes we found reference complex amplitude diffraction efficiency + for the 0th order transmission and + for the +1st order transmission.

Let us consider the difference from the reference value as an error of a tested method. Figure 4.4 presents the errors by both modifications of the FMM versus the inverse number of diffraction orders. The number of diffrac-

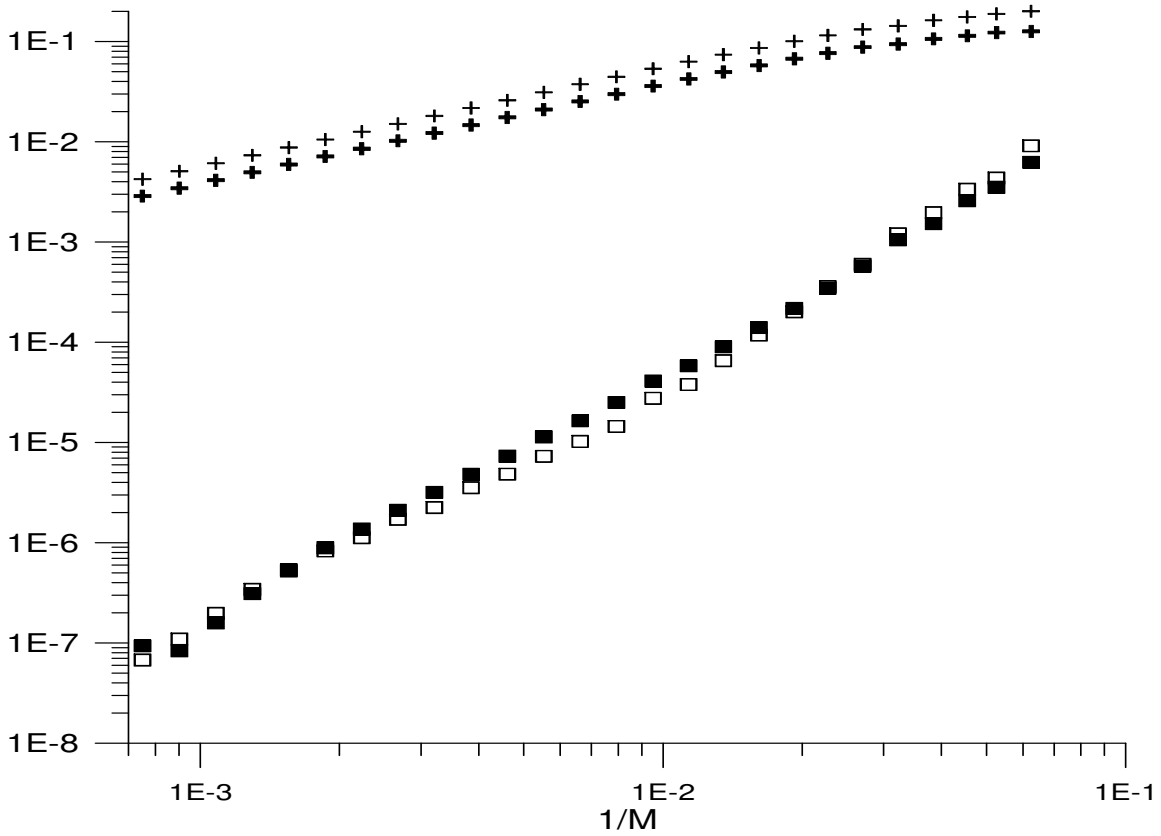


Figure 4.4: The errors versus the inverse number of modes for the original FMM (the crosses) and its modification LMT (the squares) for the 0th order (bold) and the +1st order (hollow) in transmission by the binary lamellar grating

tion orders is equal to the number of modes M that was used in calculations. One can conclude from Fig.4.4 that the original version of the FMM [3, 48] and its modification LMT [42] converge finally to the same point although the convergence of the initial version is very slow. Even for M = 1000, the diffraction efficiencies by the original FMM are far from the reference value. Figure 4.4 illustrates clearly that convergence is much better with the LMT. This shows the necessity of using such improvement for calculating diffraction efficiencies with the FMM.

4.3.2 Sinusoidal grating

The sinusoidal grating profile is a good benchmark for slicing modelling techniques for two reasons. First, the normal to the grating interface varies strongly with the slice position which strongly affects the boundary conditions. Secondly, there is an excellent reference given by the C method [1]. Applying different techniques to such structure most evidently shows difference in their behaviour. We investigated convergence of the 0-th and the +1-st transmission order field amplitudes for the following structure: grating period $d = \lambda = 1000nm$, full groove depth 500 nm, grating refractive index . The substrate is made of the same material as the grating. The second medium of the grating and the cover is air with . The incidence was from the air side at 45 degree with magnetic field aligned along grating lines (TM polarisation). Applying the C method delivers reference complex amplitude diffraction efficiency for the 0th order transmission and + for the +1st order transmission.

Contrary to the binary lamellar profile, in the calculation of sinusoidal gratings the number of slices N is an important parameter which has an impact on the convergence. Therefore, this is a two-dimensional convergence problem, where to find the best convergence path is not an elementary issue. We performed calculations with the novel and the standard techniques varying the number of diffraction orders M and the number of slices N taken into account. Taking the amplitudes given by the C method as the reference we build for each technique a surface of errors. The surfaces for LMT and the novel technique are depicted in Fig.4.5.

It follows from Fig.4.5 that for both techniques the optimum ratio between the number of slices N and the number of diffraction orders M is about 10 in order to achieve the best calculation accuracy. We kept this ratio in diffraction efficiency calculations by all the three techniques: the initial FMM, the LMT and the novel technique. The results are presented in Fig.4.6. Note first that no improvement is achieved by applying the LTM with respect to the original FMM. This means that the advantage of the LMT is completely lost in the stair-case approximation of the sinusoidal profile. Quite opposite,

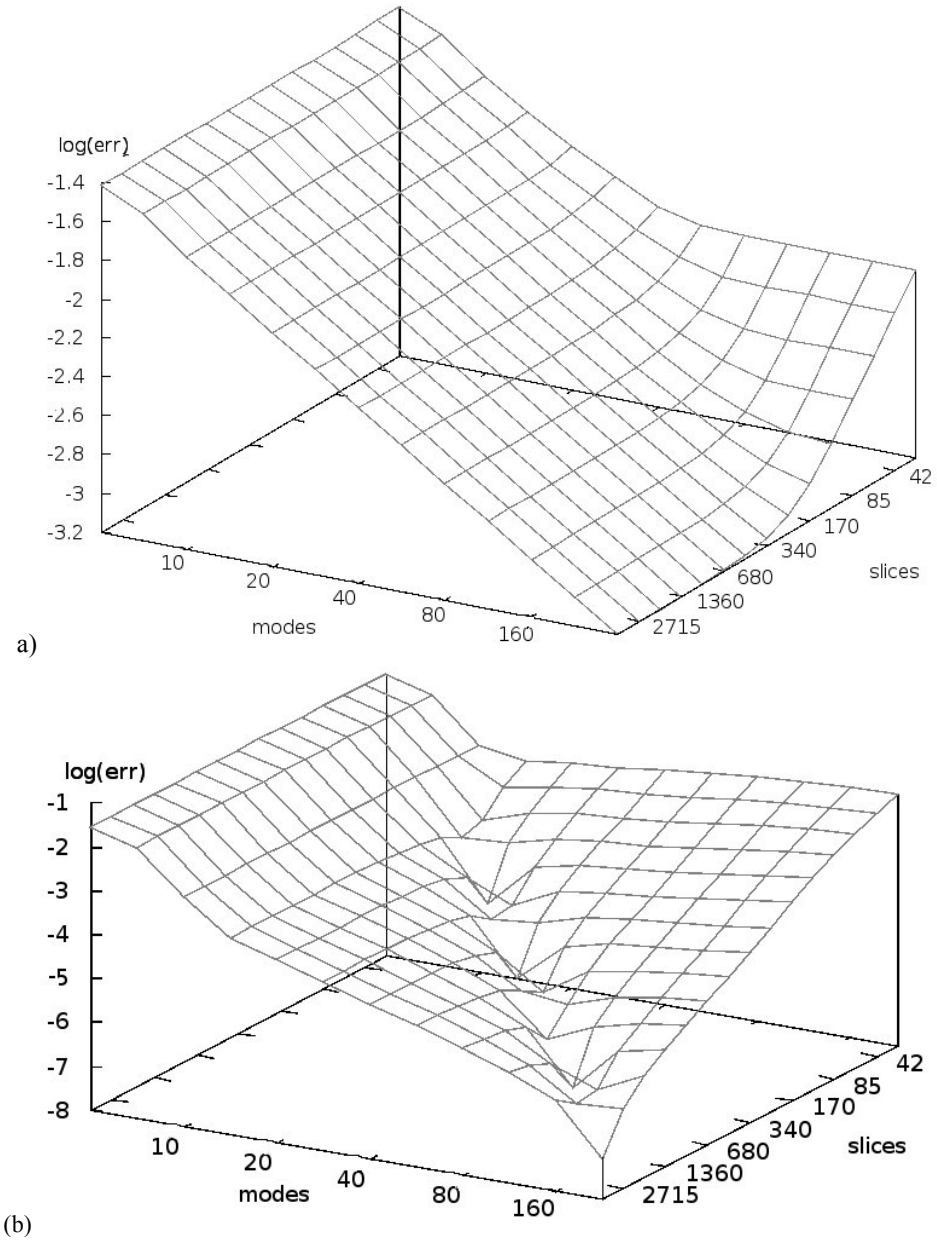


Figure 4.5: The error in calculation of the +1 order transmission of a sinusoidal grating versus number of modes M and slices N for the LMT (a) and our modification (b) of the FMM

our novel technique is relevant in the stair-case approximation and improves much the convergence of the FMM. The third example concerns the conical diffraction of a TM polarised wave on the sinusoidal grating. The structure

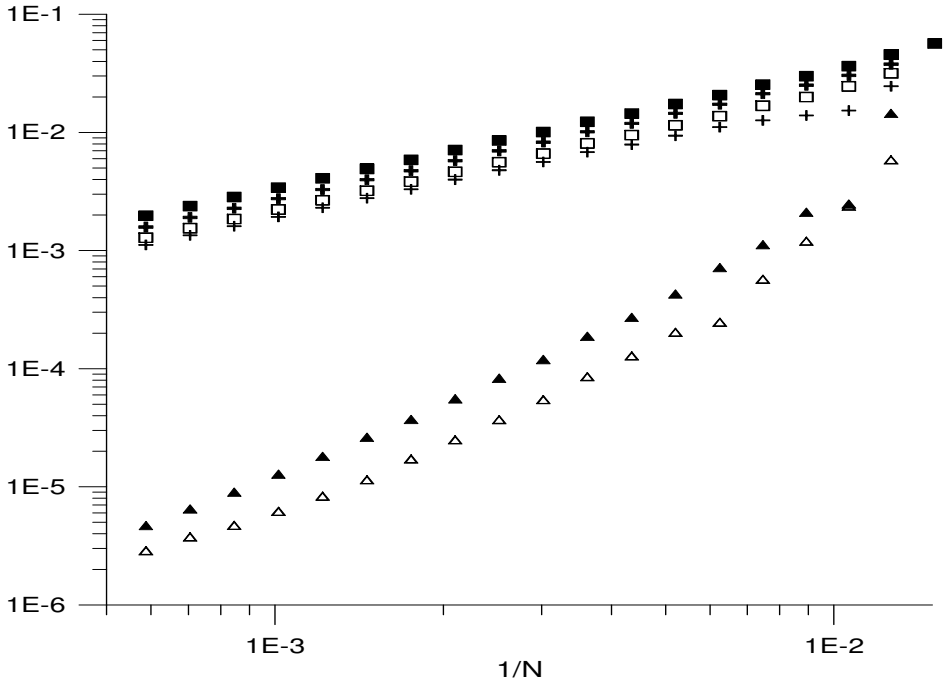
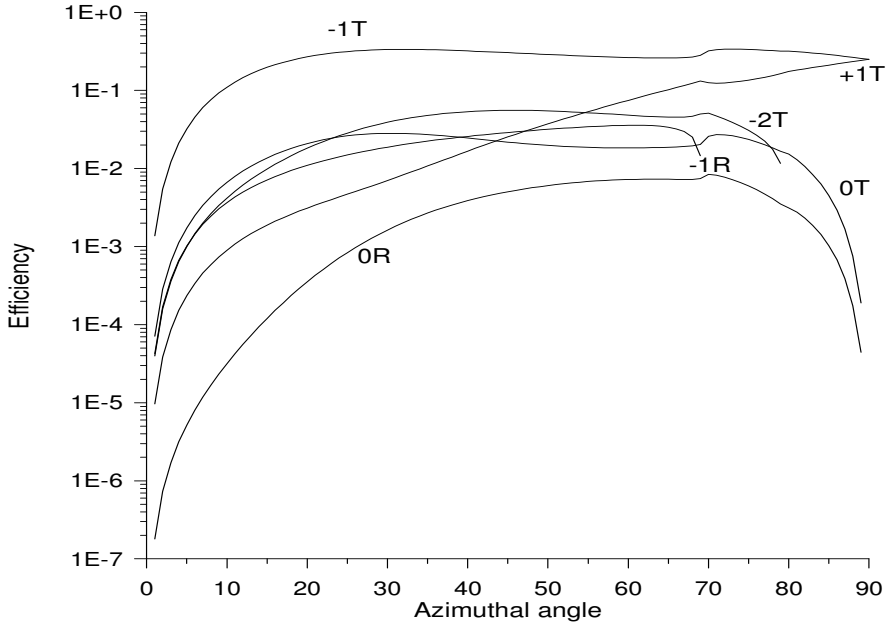
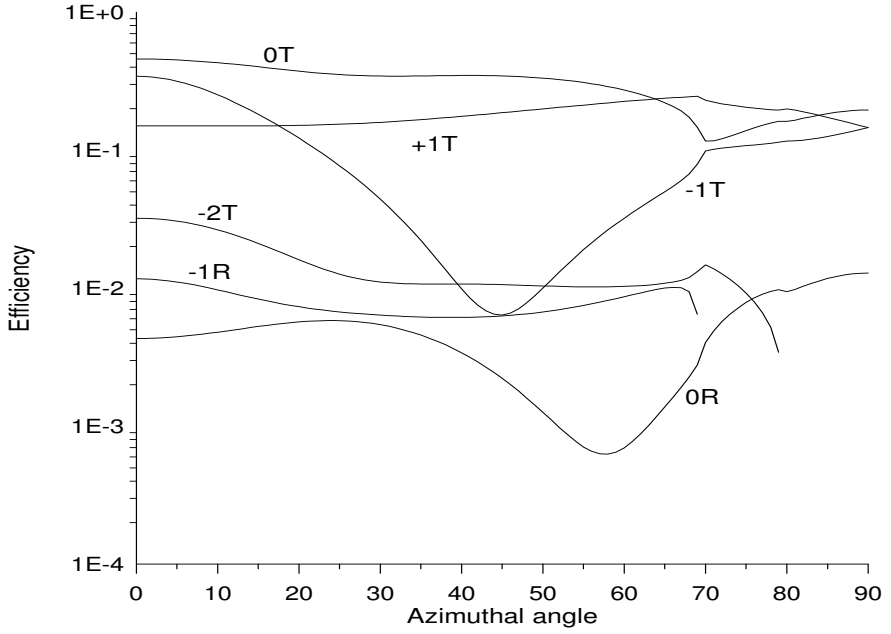


Figure 4.6: The errors versus the inverse number of slices for the original FMM (the crosses), the LMT [12] (the squares), and our technique (the triangles) for the 0th order (bold) and the +1st order (hollow) in transmission by the sinusoidal grating. The number of slices N is maintained ten times the number of modes M

is the same as in the previous example and the C method was used as a reference. Since the C method implementation to conical diffraction is not commonplace we checked also the obtained results by the Rayleigh method [39] and found coincidence between the two methods up to computer error (better than 1e-14). Calculated diffraction efficiency is given in Fig.4.7. As in the previous section the results by three different FMM techniques are compared with the reference. Figure 4.8 shows the average-square error in six propagating diffraction orders (four orders in transmission and two orders in reflection) versus azimuthal angle. The novel technique demonstrates two advantages before the other versions of the FMM. First, the accuracy and convergence of the method are improved significantly. Secondly, the energy balance test gives a good error estimation whereas in the both other versions of the FMM it rests always at the level of computer error.



(a)



(b)

Figure 4.7: Power diffraction efficiency from the sinusoidal grating in conical mount. a. TE polarised waves; b. TM polarised waves.

4.4 Conclusion

We consider this work as a next step to the improvement of the FMM. With two grating examples we demonstrated that the certain modification of the

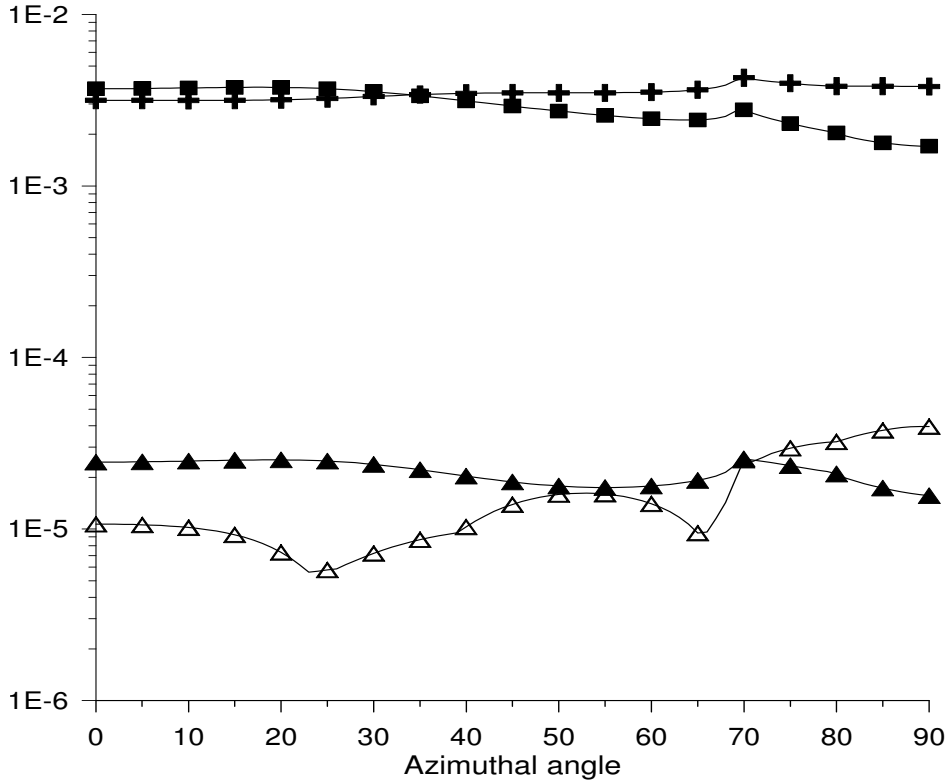


Figure 4.8: Average-square amplitude error for the original FMM (crosses), the LMT (squares), and the novel technique (filled triangles); energy balance error y the novel technique (hollow triangles)

equation system is needed which results in a sound increase in both the accuracy and the convergence of the method. One should take into account while choosing the proposed technique that considering a slanted coordinate system doubles the matrix size and increases correspondingly the computation time. Such deceleration however is well compensated by the substantial improvements offered by the novel technique.

The proposed technique is absolutely compatible with other recent improvements of the FMM such as spatial adaptive resolution [55, 56], spurious mode filtering[57], and 2D grating developments[15, 52, 53] . Therefore, it can be efficiently applied together with those techniques when calculating relief metal gratings.

Chapter 5

Building a solution for the two-dimensional grating

We can consider a two-dimensional grating as an extension of the one-dimensional case. There are additional parameters to describe a grating: periodical vector along second axis is added and permittivity distribution function $\varepsilon(x, y, z)$ becomes three-dimensional. Generally, the periodicity vectors of the grating can be non-orthogonal. We will consider gratings with orthogonal periodicity vectors but all the described steps can be adjusted to the for non-orthogonal case. Such adjustment is not given in this part because it is worth a particular research.

This chapter is split into two main steps necessary to get a solution for the diffraction problem on the two-dimensional grating: the first part deals with a search for the propagation constants and reconstructing the field distributions associated with those constants; and the second part is dedicated to a procedure of the overlap integrals calculation taken for the different field projections, construction of the transition matrices for each of the grating interfaces and the final scattering matrix con-

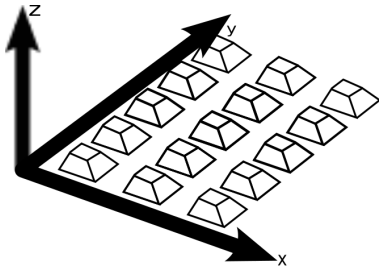


Figure 5.1: Coordinate system bounding

struction based on the transition matrices obtained in the previous stage.

We bind an orthogonal coordinate system with the two-dimensional grating with orthogonal periodicity vectors according to the following rule: $z-$axis is perpendicular to the grating surface, $x-$axis coincides with on of the grating's periodicity vectors and $y-$axis coincides with the second grating's vector, forming a right-handed coordinate system. Solutions above and under the grating we represent as a superposition of the plane waves (orders). According to the grating equation we expect a two-dimensional diffraction picture and, for each periodicity direction, we can write an equation for a wave vector projection onto the axes:

$$\vec{\mathbf{A}}_{n,m}(x,y,z,t) = \vec{\mathbf{A}}_{n,m} \exp(jk_n^x x) \exp(jk_m^y y) \exp(jk_{n,m}^z z) \exp(-j\omega t) \quad (5.1)$$

A diffracted wave thus far is characterised by two indices n and m, where n is the $x-$coordinate index and m is the $y-$coordinate index. A couple (n,m) will be also referred to as a "coordinate number". Projections of the waves on the $z-$direction in the semi-infinite media with permittivity ε are equal to:

$$k_{n,m}^z = \sqrt{\omega^2\mu\varepsilon - \left(k_n^{x\,2} + k_m^{y\,2}\right)}$$

We see, that the plane wave is propagating when the following condition is satisfied: $Re(k_{n,m}^z) > 0$ or $\omega^2\mu\varepsilon > {k_n}^{x2} + {k_m}^{y2}$. It is evident that the propagation qualities of the orders depend on both the indices (or coordinate number). $\vec{\mathbf{A}}$ in equation (5.1) expresses the complete electromagnetic vector (it can be either electrical or magnetic field as electrical and magnetic fields can be expressed in term of each other in the isotropic media case).

To simplify Maxwell's equations we use the same plane wave polarisation definition as introduced in the 2.1.1 with the only change being that now each order is characterised by its coordinate number (two indices). The TE-polarised wave is denoted $E_{n,m}$ and the TM-polarised wave $H_{n,m}$. Defining plane wave vector projection on the grating surface like: $\gamma_{n,m} = \sqrt{k_n^{x\,2} + k_m^{y\,2}}$, indexed by coordinate number, we may write plane waves decomposition with

using coordinate number :

$$
\begin{aligned}
E_x &= \sum_{n,m} \quad \tfrac{k_m^y}{\gamma_{n,m}} E_{n,m} \quad -\frac{1}{\omega\varepsilon}\frac{k_{n,m}^z k_n^x}{\gamma_{n,m}} H_{n,m} \\
E_y &= \sum_{n,m} \quad -\tfrac{k_n^x}{\gamma_{n,m}} E_{n,m} \quad -\frac{1}{\omega\varepsilon}\frac{k_m^y k_{n,m}^z}{\gamma_{n,m}} H_{n,m} \\
E_z &= \sum_{n,m} \quad 0E_{n,m} \quad -\frac{\gamma_{n,m}}{\omega\varepsilon} H_{n,m} \\
H_x &= \sum_{n,m} \quad \tfrac{1}{\omega\mu}\tfrac{k_n^x k_{n,m}^z}{\gamma_{n,m}} E_{n,m} \quad +\frac{k_m^y}{\gamma_{n,m}} H_{n,m} \\
H_y &= \sum_{n,m} \quad \tfrac{1}{\omega\mu}\tfrac{k_m^y k_{n,m}^z}{\gamma_{n,m}} E_{n,m} \quad -\frac{k_n^x}{\gamma_{n,m}} H_{n,m} \\
H_z &= \sum_{n,m} \quad -\tfrac{\gamma_{n,m}}{\omega\mu} E_{n,m} \quad +0H_{n,m}
\end{aligned} \tag{5.2}
$$

where axes dependencies $\exp(jk_n^x x)\cdot\exp(jk_m^y y)\cdot\exp(\pm jk_{n,m}^z z)$ are eliminated. A plane wave decomposition allows us to use superposition of the TE- and TM-polarised modes noted with indices (n, m) due to the orthogonality of the plane-wave functions.

5.1 Propagation constants search procedure. Reconstruction of the modal fields

5.1.1 Requirements for the modal functions and imposed conditions

The diffraction problem on a two-dimensional grating is a quite non-trivial task. Nowadays, existing methods are capable to provide some estimation on the diffraction efficiency in particular structures. But the accuracy of the results obtained for the metal gratings is very poor and some more method is necessary for this type of gratings. Even though there are methods developed for the calculation diffraction efficiency on the two-dimensional periodical structures, there are no publications concerning solution with the true-modal

method.

A modal approach is based on a definition of the modal basis, on a selection of the modal functions, resolution of a general problem and subsequent calculation of the amplitudes satisfying the particular problem. Thus far, first aim for a resolution of the problem within modal method is to define the mode and modal composition which will be used to describe solution.

We introduce a mode as a particular Maxwell's equations system solution, possessing permanent wave vector projection along the depth-coordinate ($z-$axis in the geometry described above). We also refer to the combination of the electrical and magnetic fields distribution and corresponding to the distribution propagation constant, as "mode". So, mode is defined propagation constant value (modal constant) plus modal fields' distributions corresponding to mode, propagating upwards or downwards (distributions depend on a propagation direction in general case).

5.1.2 The basic grating structure for modal method implementation

The modal solution is depth-independent, we assume that the profile of the grating is also depth-independent. So the permittivity distribution function is thus function of two coordinates and remains the same for any depth: $\varepsilon(x, y, z) = \varepsilon(x, y)$ In case of a grating profile, where this distribution function is depth-dependent,a slicing technique should be applied reducing the problem to several sub-problems appropriate for the modal method.

5.1.3 Problem formulation

Now we will consider criteria that are applied to a representation of the modes and will give a clue to construct a modal basis. A grating periodicity in two directions allows us to define the distribution of a mode over an elementary structure cell. At the same time, from the conditions imposed by the field behaviour of the incident wave, we get modal fields matching conditions on the boundaries of the two neighbouring elementary cells (and displaced on an

distances equal to the product of the integer by the period along the axis).

After the bounding of the coordinate system with the grating we can attribute coordinate indices to the periods of the grating, so that the period along the $x-$axis is denoted Λ_x and along the $y-$axis as Λ_y. We illuminate the grating by a plane wave with wavelength λ at an azimuthal angle Θ (taken from the $z-$axis direction) and polar angle ϕ taken from the $x-$axis (see fig.5.2). Diffracted and reflected wave vector projections are defined by the following relations:

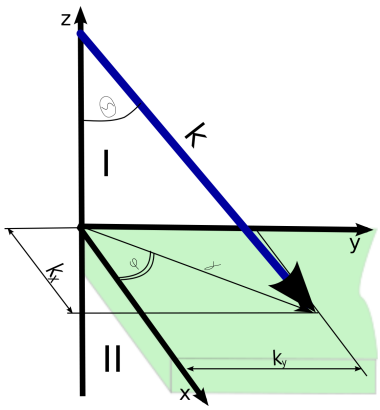


Figure 5.2: Incidence definition

$$\begin{aligned} k_n^x &= k_0^x + \tfrac{2\pi}{\Lambda_x} n \\ k_m^y &= k_0^y + \tfrac{2\pi}{\Lambda_y} m \\ {k_{n,m}^{z,i}}^2 &= \omega^2 \mu \varepsilon_i - {k_n^x}^2 - {k_m^y}^2 \end{aligned}$$

where ε_i is the permittivity of the semi-infinite media i, $k_{n,m}^{z,i}$ is the $z-$ projection of the wave vector in the media i, (n, m) is the coordinate number described earlier. The components of $\vec{k_0}$ are defined as:

$$\begin{aligned} k_0^x &= \frac{2\pi}{\lambda} \sqrt{\varepsilon_I} \sin(\Theta) \cos(\phi) \\ k_0^y &= \frac{2\pi}{\lambda} \sqrt{\varepsilon_I} \sin(\Theta) \sin(\phi) \end{aligned}$$

The periodical conditions imposed by the incident wave lead to the equalities:

$$\begin{aligned} \vec{E}(x + u \cdot \Lambda_x, y + v \cdot \Lambda_y) &= \vec{E}(x, y) \times \exp(j k_0^x \cdot \Lambda_x u) \exp(j k_0^y \cdot \Lambda_y v) \\ \vec{H}(x + u \cdot \Lambda_x, y + v \cdot \Lambda_y) &= \vec{H}(x, y) \times \exp(j k_0^x \cdot \Lambda_x u) \exp(j k_0^y \cdot \Lambda_y v) \end{aligned}$$

The same conditions should be satisfied by the modal fields distributions propagating inside the grating region. The periodicity conditions are common for all the modes.

5.1.4 Bloch's modes and orthogonality

According to Bloch's theory, if we have a problem with periodical operator, it is possible to represent all the solutions in the form $\exp(j\vec{k}\vec{r}) \cdot \vec{u}_k(\vec{r})$. It is possible to pick out the exponential term in some specific direction and represent any problem solution as a superposition of such solutions. Bloch's theorem is applicable for the problems where a linear operator can be written. From this theorem a possibility arises to construct a modal basis for two-dimensional diffraction problem, because boundary conditions are periodical in two directions (periodical condition), Maxwell's equations is system of linear operators on the field components and it's possible to construct linear operator of the problem (existence of the problem linear operator). Furthermore, it is possible to build an orthogonal basis with the functions in the form $\psi_c(x, y) \cdot \exp(\pm\beta z)$ for each field component c. We call as a modal representation of the diffraction problem solution a set of functions for field components E_x, E_y, E_z, H_x, H_y, H_z in the form $f(x, y) \cdot \exp(\pm\beta z)$, satisfying Maxwell's equations. In one-dimensional true-modal method there is only one function corresponding to each propagation constant. Fields' projections of the mode are reconstructed by linear operators of a fixed form. Functions $f(x, y)$ can be arbitrary in general case and differ for different propagation constants. We will show one approach to construct such functions.

5.1.5 Product-like modal function suggestion

A form of the base function in the one-dimensional grating case is dictated by a translation symmetry of the grating in one of the directions. The translation symmetry allows to reduce a search of a modal functional representation to the construction of the one-dimensional function (a coordinate dependence along the translation axis is pre-defined by incident plane wave). The one dimensional function set, where each function satisfies Maxwell's equations, is limited by plane waves and their derivatives. In the two-dimensional case of the field and its projections distribution there is much more freedom to select an appropriate functional basis. On the other hand, this freedom does not bring simplification to the modal basis construction because solutions

satisfy Maxwell's equations and are mutually orthogonal. The periodicity of the field solutions allows to represent projection on the axis p of the modal fields F as the sum

$$F_\beta^p(x,y) = \sum_{i=0}^{\infty} f_i^\beta(x) \cdot g_i^\beta(y)$$

This is an infinite sum of the products of the single argument functions. The $g(y)$ and $f(x)$ functions can have arbitrary form for different values β of the propagation constant. Such functional behaviour allows to look for the function in the form of a linear combination of the exponential functions. There should be generating basis common for all the projections. We denote propagation constant β and infinite vector $(a_i\psi_i(x)\rho_i(y))$ where each product contain amplitude a_i and axis dependencies $\psi_i(x)$, $\rho_i(y)$ of some form. Introducing vector-strings $\left(\mathcal{L}_j^p\right)$ of the linear differential operators, so that for each projection j of the field p will be result of the operator action:

$$F_\beta^p(x,y) = (\mathcal{L}_l^p(\beta)\cdots) \times \begin{pmatrix} a_k\psi_k(x)\rho_k(y) \\ \vdots \end{pmatrix} = \sum_{i=0}^{\infty} f_i^{\beta,p}(x) \cdot g_i^{\beta,p}(y)$$

we reach the aim that each field projection is an infinite sum of the separating functions and generating vector exists for each mode. Fixation of the function form leads to the resolution of the representation uniqueness problem.

As an example of this approach we consider the most simple structure for the modal method: a rectangular elementary cell made of a material with permittivity ε_2 with another rectangular of the material ε_1 inside. The grating periods are Λ_x and Λ_y. Sizes of the elementary cell are the same. Internal rectangular has measurements l_x and l_y along the corresponding axes. We will use the following notation $mr_x = \frac{l_x}{\Lambda_x}$ and $mr_y = \frac{l_y}{\Lambda_y}$. These values correspond to the filling factor for the one-dimensional grating. In the two-dimensional grating there are two variables to describe the filling along each of the axes. But it is confusing to call these values filling factors since actually the filling factor is a product of these values, i.e. the ratio of the area with material ε_1 to the all elementary cell area.

We place the coordinate system so that the left-bottom corner of the in-

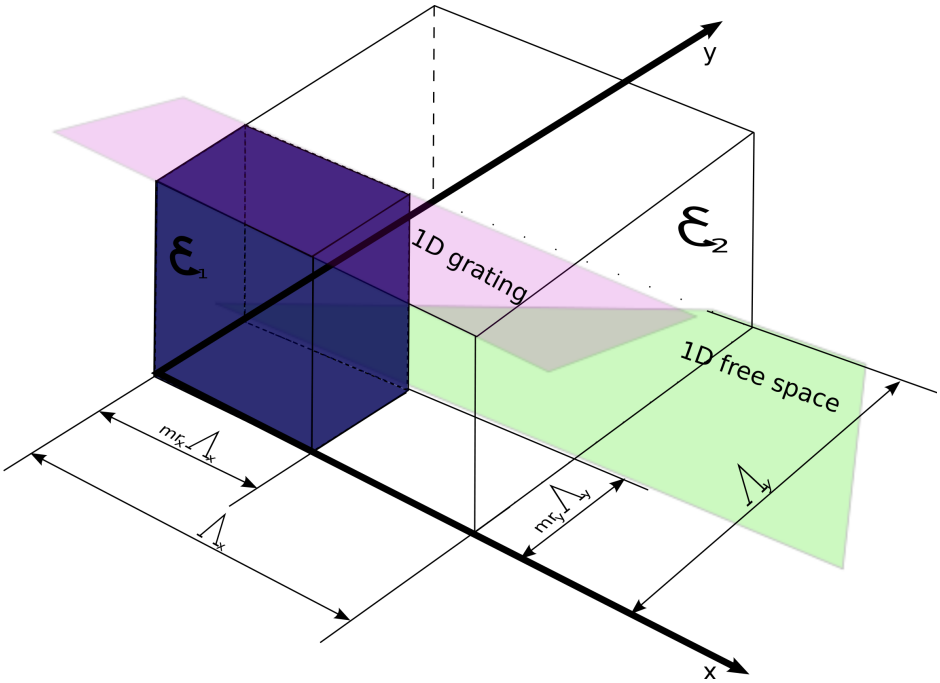


Figure 5.3: Elementary cell notation

ternal rectangular coincides with the zero-point and the grating periodicity vectors coincide with the $x-$ and $y-$ axes. The elementary cell can this way be divided into two sub-cells: one of them reminds one-dimensional elementary structure and another one is volume with constant permittivity. From the one-dimensional case we know that Rayleigh's plane wave decomposition of the field is most suitable for regions with constant permittivity. We will thus represent the fields inside that region as a superposition of the plane waves. For the region with several materials we will use the true-modal field representation. Such a field representation allows to reduce the procedure of the function search to the calculation of the modal and order's amplitudes. Plane wave basis and modal basis in conical mount are complete basis. It follows that any field inside the introduced elementary cell can be represented in these basis and the representation will be unique and complete.

Modal field decomposition and order field decomposition allow to express any field projection. Therefore, we transform the problem to constructing a function to the search for the amplitudes for the already known basis.

5.1.6 More complex profiles representation

The true modal method is appropriate for the structure consisting of alternating slabs per a period. For more complicated profiles a slicing technique should be used. Extended Fourier-modal method allows to obtain field of the complicated problem. Choice of each method is determined by the profile and accuracy of the field reconstruction difficulty. The true- and Fourier- modal methods were discussed in detail in the previous chapters. Each method can be used to obtain propagation constant values and to reconstruct fields. These modal methods provide uniform field representation and guarantee uniqueness of the solution.

5.1.7 Construction of the dispersion function

The periodicity condition should be fulfilled by the complete problem solution and by each modal function. To construct the dispersion equation (an equation which defines the modal propagation constants) we follow the steps described in the following. First step is to select a one-dimensional periodic structure, so that repetition of that structure gives us original structure. We consider a structure which repeats $x-$distribution of the two-dimensional structure and equivalent to one period along $y-$direction. This structure is infinite in the $z-$ and $x-$ directions and finite along $y-$direction. This extract will be also referred as "elementary stripe" of the original structure. Solution will be constructed for configuration when a structure is illuminated with waves, so that waves' vector projection on the $z-$axis can be arbitrary complex value (even pure imaginary), projection on the $x-$axis coincides with the k^x projection of the

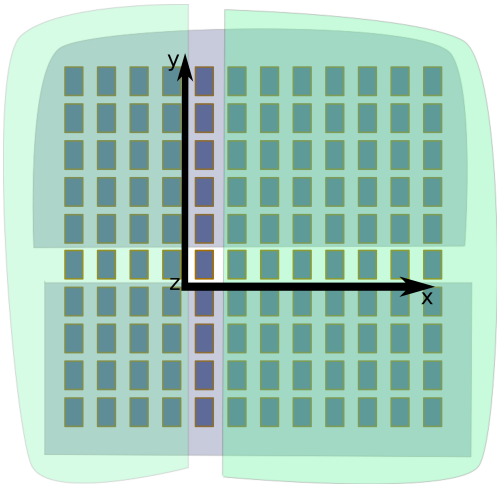


Figure 5.4: Top view on the grating

original problem. The one-dimensional problem in different coordinate system is well studied. This configuration corresponds to the conical mount with the only prerequirement that wave vector projection along the translational axis can be imaginary (or any complex value). It was mentioned that the true-modal solution for the one-dimensional structure is not restricted to real components of the wave vector projection along the axis with the translational symmetry. Therefore, the problem geometry can by resolved with already existing method. Depending on the complexity of the extracted structure we can use true-modal or extended Fourier-modal method to construct scattering matrix of the stripe. The choice of the method to be used is determined by the structure geometry, required precision and computational time.

The resulting scattering matrix gives the relation between the waves, incoming onto the elementary stripe waves and the outgoing waves. One can construct the scattering matrix of an elementary stripe corresponding to some k^z projection (modal propagation constant value), associating incoming and outgoing waves (fig. 5.5):

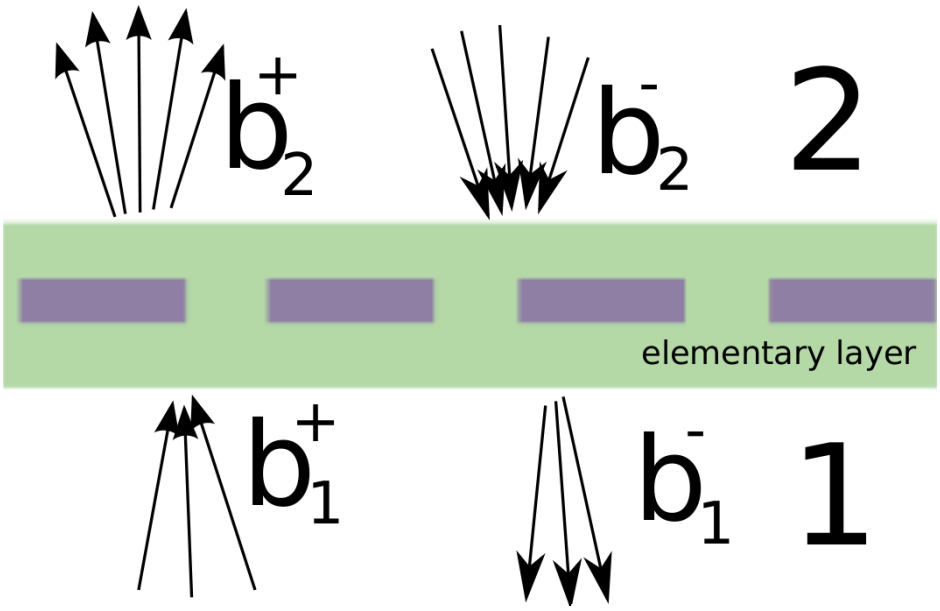


Figure 5.5: Definition of incoming and outbound waves (orders)

$$\begin{pmatrix} b_2^+ \\ b_1^- \end{pmatrix} = S \begin{pmatrix} b_2^- \\ b_1^+ \end{pmatrix}$$

Now we return to the original problem. From the periodical conditions imposed by an incident wave and its wave vector projection we can conclude, that equation:

$$\begin{pmatrix} b_2^+ \\ b_2^- \end{pmatrix} = \exp(j^i k^y \Lambda_y) \begin{pmatrix} b_1^+ \\ b_1^- \end{pmatrix}$$

is valid for all the components. Introducing the diagonal matrix Γ with diagonal elements equal $\exp(j\ ^i k^y \Lambda_y)$ we can write previous equation with some rearrangement of the components:

$$\begin{pmatrix} b_2^+ \\ b_1^- \end{pmatrix} = \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma^{-1} \end{pmatrix} \begin{pmatrix} b_1^+ \\ b_2^- \end{pmatrix} \tag{5.3}$$

Rearranging the scattering matrix $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ to the form $\hat{S} = \begin{pmatrix} B & A \\ D & C \end{pmatrix}$ gives

$$\begin{pmatrix} b_2^+ \\ b_1^- \end{pmatrix} = \hat{S} \begin{pmatrix} b_1^+ \\ b_2^- \end{pmatrix} \tag{5.4}$$

At the same time with validity (5.3) we can subtract from (5.4) and get the matrix equation:

$$\left(\hat{S} - \hat{\Gamma}\right) \begin{pmatrix} b_1^+ \\ b_2^- \end{pmatrix} = 0 = \lambda_m \cdot \begin{pmatrix} b_1^+ \\ b_2^- \end{pmatrix} \tag{5.5}$$

where $\hat{\Gamma} = \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma^{-1} \end{pmatrix}$. This equation has non-trivial solution when determinant of the matrix $\left(\hat{S} - \hat{\Gamma}\right)$ is equal zero:

$$det(\hat{S} - \hat{\Gamma}) = 0$$

Taking into account that matrix $(\hat{S})$ is reordered matrix S which is actually

the function of propagation constant β, we introduce function $d(\beta)$:

$$d(\beta) = det\left(\hat{S}(\beta) - \hat{\Gamma}\right) \tag{5.6}$$

The condition $d(\beta) = 0$ is verified when $\begin{pmatrix} b_1^+ \\ b_2^- \end{pmatrix}$ is eigenvector of the $\left[\hat{S} - \hat{\Gamma}\right]$ matrix corresponding to zero eigenvalue. On the other hand, this eigenvector satisfies the periodicity condition along the $y-$axis.

The constructed function $d(\beta)$ is the dispersion equation of the two-dimensional problem and resolution of this equation gives the propagation constant values of two-dimensional grating modes. This function maps complex argument to the complex result $d(\mathbb{C}) \to \mathbb{C}$. It is difficult to use traditional properties of the function for analysis or to apply the functional analysis methods to this function. This function is defined numerically and can be analysed only on the basis of the calculated values.

The propagation constant search procedure differs completely from the diagonalising of the matrix in the RCWA method. In the method described above, whatever method provides scattering matrix, the dispersion function is constructed. And propagation constants correspond to zero values (roots) of the dispersion function.

5.1.8 Eigenvector amplitudes and field reconstruction

Having the propagation constant (β value satisfying the dispersion equation) the problem of constructing (retrieving) vector of amplitudes is not difficult from the mathematical point of view: this is a standard procedure of finding the eigenvector of the given matrix. Depending on the used method to construct the scattering matrix, several eigenvectors of the matrix and the corresponding set of vector amplitudes can be found. From the entire set of the vectors we need to allocate one eigenvalue which coincides with 0, because look for a vector corresponding to the equation:

$$\left(\hat{S} - \hat{\Gamma}\right)\begin{pmatrix} b_1^+ \\ b_2^- \end{pmatrix} = 0 \cdot \begin{pmatrix} b_1^+ \\ b_2^- \end{pmatrix}$$

Having set of the field components amplitudes defined with a predetermined rule, the reconstruction of the field distribution over the elementary cell is reduced to the summation of the components, multiplied by the corresponding phase factors which depend on the wave vector of each component. The matrix eigenvectors are defined up to an arbitrary factor. This allows us to scale the amplitude of each component to meet the criteria of a mode normalisation. As the most convenient normalisation criterion we find the following:

$$\langle P_z \rangle = 1$$

where we use $\vec{P} = \frac{[\vec{E},\overline{\vec{H}}^*]}{2}$ vector product of the electrical field of the constructed mode with magnetic field of the congruent problem formulation. This criterion is applicable for both propagating and decaying modes. The normalisation allows to compare the components of the fields attributed to different propagation constants. Without the normalisation we still can compare field components of the particular mode. but also between modes with different propagation constants.

5.1.9 Realisation example and results

To verify the approach we have chosen the most simple structure satisfying the basic requirements of the modal method applicability. The elementary cell consists of a rectangle with sides $\Lambda_x = 300nm$ $\Lambda_y = 400nm$ of a material with a reflection coefficient $n_2 = 1$, another rectangle is located in the interior of the cell (reflection index is $n_1 = 2.2$ in dielectric case and $\varepsilon = -5$ in metal case) with $mr_x = mr_y = 0.5$ (sides are equal to $mr_x \cdot \Lambda_x$ and $mr_y \cdot \Lambda_y$ respectively). Considered structure is periodic in the $x-$ and $y-$ directions and infinite in the $z-$direction. Because we only interested in the propagation constants values and field projections distributions with respect to the period, the depth of the structure is not defined at this stage. We use the same condition for our infinite-depth structure as they are the same as they are for the finite-depth structure illuminated with plane wave with wavelength equal $\lambda = 632.8nm$.

For the selected structure, we have free choice, relative direction along

which the infinite elementary stripe will be taken. If such choice is arbitrary, one expects it to be independent of the selected direction and the equality of the obtained propagation constants for both possible selections. This property can be used to verify the validity of the applied method and such equality proves that dispersion function is correct. As it was mentioned earlier we construct elementary stripe so it repeats grating in the $x-$direction. We have an infinite in the $x-$direction elementary stripe and we will construct the scattering matrix of this stripe. The elementary stripe consists of two sub-stripes: there is a sub-stripe with a refractive index n_1 above a one-dimensional grating, such that the period is Λ_x, the depth of the grating equals $mr_y \cdot \Lambda_y$, filling factor of the grating is mr_x. The scattering matrix for a one-dimensional grating with rectangular walls can be obtained most conveniently with the assistance of the RCWA or True-Modal methods (both are well-developed for one-dimensional gratings). Scattering matrix of the stripe of a constant permittivity can be written analytically. Components of this matrix are easy to obtain and this does not create any problem. Multiplication of two scattering matrices according to the rule of multiplication of S matrices gives the resulting scattering matrix of an elementary stripe.

At this stage, it is necessary to take into account that the way of splitting the elementary stripe into substripes is a free choice and it should be done in regard of the numerical stability and the calculation error reduction. We believe that for the structure under consideration the smallest error is obtained when we represent the elementary stripe in the form of a stripe with constant permittivity of the half-height plus a stripe that represents grating and after another half-stripe with a constant permittivity. We shifted zero point downwards by the half-stripe depth. We have not introduced any changes into the problem because according to the periodic conditions the periodicity requirement remains the same. From the numerical point of view, we shifted into the zone where the amplitudes of the evanescent plane waves (orders) are of the processor zero and do not introduce strong perturbations of the solution. Scattering matrix of each sub-stripe depends on the wave vector in the $z-$direction and the resulting matrix depends on tested as a solution value β. We choose the tested value based on some root-search procedure,

build the scattering matrix $S(\beta)$ of the elementary stripe, calculate value of the function $d(\beta)$ (eq.5.6) to test if chosen value of β gives zero function value. This operation is repeated until necessary number of the propagation values is found.

In theoretical development of the method the scattering matrix is infinite while numerical modelling requires consideration of the matrices with the finite size. The actual size of the matrix is limited by the time required to complete the calculations, by the accumulation of errors arising from numerous numerical operations and amount of available memory. The lower bound of the matrix size is dictated primarily by the accuracy which is aimed to be achieved when using the method. From the numerical application point of view, when we construct a matrix of the elementary stripe and test values to satisfy dispersion equation, new free to chose parameter appears: the number of modes used while finite representation or number of variables describing the field function. The accuracy as a function of representation number can be observed on the fig.5.6 for the cases of metal and dielectric gratings with the parameters provided above and for the case of the light incident at $\Theta = 15$ and $\phi = 15$ degrees. The values of the first propagation constants for these

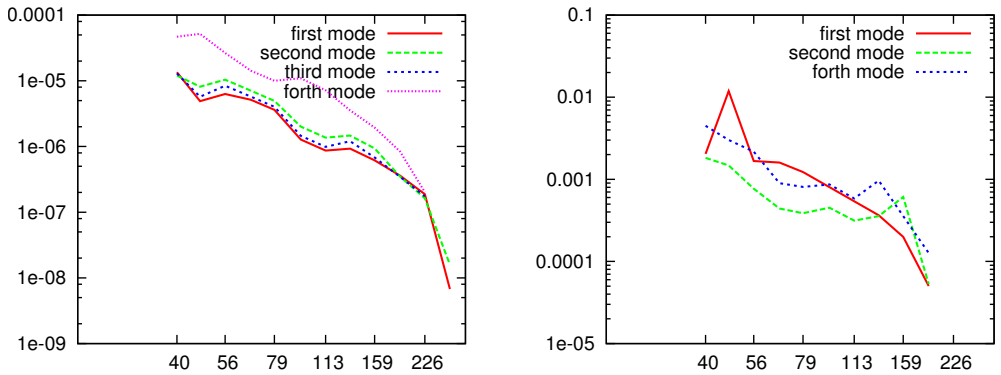


Figure 5.6: The difference between the propagation value obtain with number of modes $n = 269$ and the value obtained with corresponding number of modes for dielectric(a) and metal (b) cases.

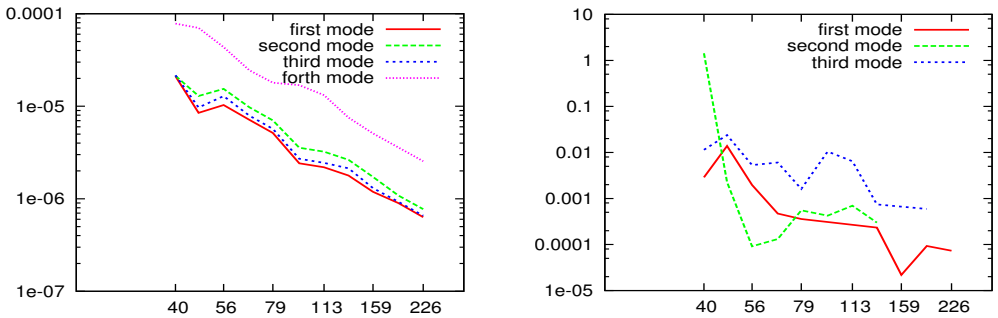
cases are provided in table 5.1.

Pro. Number	Value	
	Dielectric	Metal
1	1.573839	1.57886
2	1.37209	1.429
3	0.55449	0.8083
4	$0.02567 + i0.736553$	$i0.812$
5	$-0.02567 + i0.736553$	$i1.0441$
6	$0.12604 + i1.001301$	$i1.1612$
7	$-0.12604 + i1.001301$	$i1.7817$
8	$i1.27897$	$0.15095 + i2.2278$

Table 5.1: First propagation constant values

As it was mentioned earlier, as a criterion for evaluation of the results reliability for the two-dimensional grating the results for the rotated original grating can serve and equality of the propagation constants obtained for two configurations. After replacement of the coordinate system $\begin{matrix} x \to y \\ y \to -x \end{matrix}$ and changing the angle $\phi \to 90 - \phi$ the modes get the same propagation constants. The difference of the calculated propagation constants between original and rotated problem as a function of the representation number is shown on the fig.5.7.

We can see that in metal and dielectric cases, the values of the propaga-



a) b)

Figure 5.7: Difference between propagation constants of the original and rotated problem as a function of the representation number. Dielectric case (a) and metal case (b)

tion constants converge to each other and we can conclude that the method

Peng et al. [49]	Moharam et al. [58]	Lalanne [59]			Li [60]	Schuster et al [52]	True modal method
		0.0	0.5	1.0			
1.7790	1.8019	1.7806	1.7913	1.8003	1.7901	1.7903	1.7903
1.5270	1.5416	1.5395	1.5326	1.5235	1.5348	1.5278	1.5277
0.2691	0.2933	0.2906	0.2919	0.2927	0.2921	0.2912	0.29117
-0.2105	-0.2150	-0.2246	-0.2297	-0.2369	-0.23498	-0.2267	-0.2268
-0.8106	-0.7475	-0.74438	-0.74098	-0.7361	-0.7406	-0.7397	-0.7396
-1.3819 - i*0.202	-1.3568 - i*0.2219	-1.3585 - i*0.2379	-1.3587 - i*0.2335	-1.3581 - i*0.2316	-1.364 - i*0.2336	-1.3524 - i*0.2365	-1.3522 -i*0.2368
-1.3819 +i*0.202	-1.3568 +i*0.2219	-1.3585 +i*0.2379	-1.3587 +i*0.2335	-1.3581 +i*0.2316	-1.364 +i*0.2336	-1.3524 +i*0.2365	-1.3522 +i*0.2368

Table 5.2: Comparison of the propagation constant values obtained with different RCWA method modification and true modal method

yields reliable values of the propagation constant including the metal grating case. It is interesting to compare the values obtained by different methods for the same structure without making any changes or modifications to the structure (rotation, slicing). For a considered structure RCWA is suitable method. Propagation constant values in case of the RCWA method depend on the total product of the two values: decomposition number along one and perpendicular directions.

Values obtained by the RCWA in the dielectric grating case coincide up to a good accuracy with the values obtained with the true-modal method. For the benchmark the same structure was used, but incidence was taken at $\Theta = 30$ and $\phi = 40$ degrees. The obtained results are gathered in table 5.2. The developed method shows that propagation constants differ from each other and also differ from the true-modal method results. Thus far, true modal method is already used as a benchmark for the existing and widely used RCWA methods. We can also see that coupled modes are found by all the methods and can conclude that these modes are not erroneous.

In the metallic structures case RCWA method creates spurious modes. These modes are not solutions of the diffraction problem. They are solutions of the reciprocal problem in Fourier space domain.

After we get the eigenvector corresponding to the zero eigenvalue of the $[\hat{S} - \hat{\Gamma}]$ matrix, we have the vector of the amplitudes of the field functions at the boundaries of the elementary stripe. The restoration of the field in the region of constant permittivity is performed by multiplying on the phase coefficients, determined by the projection of plane waves on the coordinates. To restore the field in the region with non-constant permittivity when the true-modal method is used to represent the field inside that region it is necessary to restore the modal amplitudes which correspond to the set of plane wave amplitudes (in the region(s) of constant permittivity). It can be achieved by multiplying the sets of amplitudes taken at the boundary by the transition matrix of the boundary. The restoration of the fields inside the grating consists of the summation of each mode contribution. We can write with the generalisation purposes that in both regions the field will be:

$$\begin{aligned} E_{pr}(x,y) &= \sum_{i=0}^{MAX} L_{pr}^{ee}\left(\Psi_i^e(x,y)\right) + L_{pr}^{he}\left(\Psi_i^h(x,y)\right) \\ H_{pr}(x,y) &= \sum_{i=0}^{MAX} L_{pr}^{eh}\left(\Psi_i^e(x,y)\right) + L_{pr}^{hh}\left(\Psi_i^h(x,y)\right) \end{aligned} \tag{5.7}$$

where the subscript pr denotes the the projection, L_{pr}^{ft} is an operator which returns field f contribution into the field t of the projection pr. Some of these operators can be zero and the form of the operators is different for plane waves and the modal representation of one-dimensional grating. As a result of such recovery operation we get six field distributions: there are three projection of two different fields. Each two-dimensional mode is characterised with a set of vector amplitudes and propagation constant. Vector of amplitudes generates all related field projections.

The normal incidence is most suitable to check the field restoration method. In the case of normal incidence there is no need to take into account phase shift with respect to the incident wave. The periodic conditions at normal incidence are:

$$\begin{aligned} \vec{F}(x+\Lambda_x, y) &= \exp(j\dot{\imath}k^x\Lambda_x)\vec{F}(x,y) = 1\cdot\vec{F}(x,y) \\ \vec{F}(x, y+\Lambda_y) &= \exp(j\dot{\imath}k^y\Lambda_y)\vec{F}(x,y) = 1\cdot\vec{F}(x,y) \end{aligned} \tag{5.8}$$

The field repeats itself over a period in both the periodicity directions and restored field should coincide at the boundaries of the elementary cell. A field projection distributions for different modes for the dielectric structure defined earlier are presented in the fig.5.8-5.10.

All figures are suitable for the field distribution analysis with the help of the visualising programs which allow to rotate the figures and see distribution in different projections. The values along the axes are eliminated as they do not provide any useful information. The modal fields are not normalised and comparison of the absolute values can not be done. The teeth with the refractive index $n = 2.2$ is centred with respect to each figure. The bounce of the fields is caused by the small representation number (40 was used).

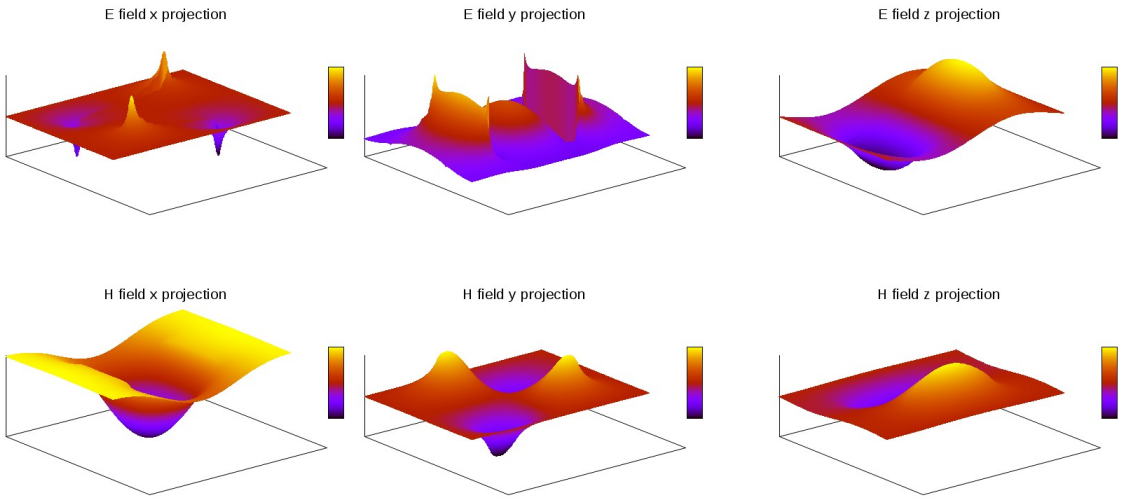


Figure 5.8: Three dimensional representation of the first mode fields projections in dielectric case

On the fig. 5.8 we can see that the distribution of the E_y and H_x does not intersect the zero plane. The E_y field is always positive and H_x field is negative everywhere. We can find similar property of the first TE-mode in one-dimensional case. On the distributions of the E_x and H_y field we can see strong anti-symmetrical disturbance of the field on the corners of the teeth. Thus far, all the field is concentrated on the corners. We can also see that

the E_x field is discontinuous along $x-$ direction (the intensity of the peak is higher outside the teeth) but is continuous everywhere along $y-$direction.

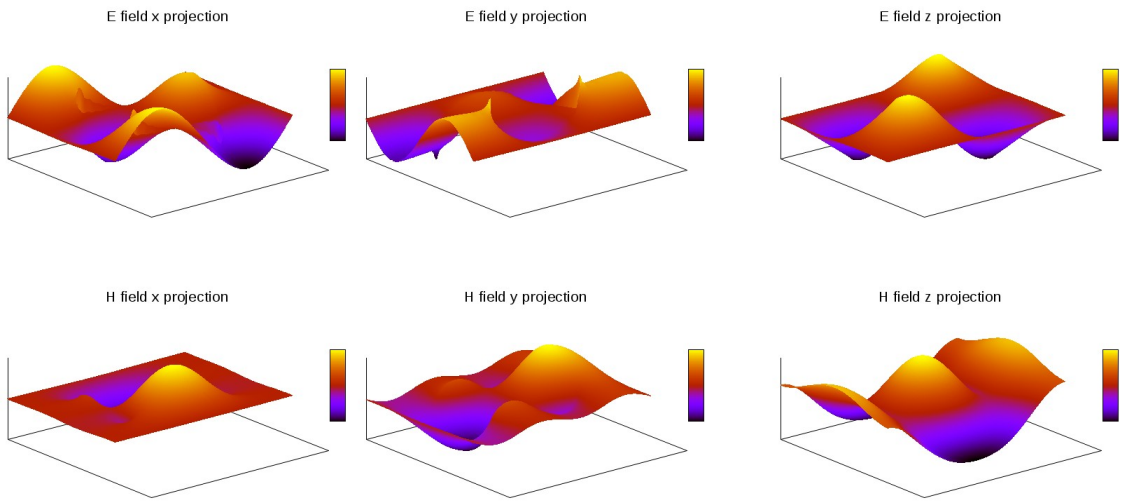


Figure 5.9: Three dimensional representation of the 10-th mode fields projections for dielectric case

On the fig.5.9 we can see that each field projection intersects the zero plane. We still can find symmetrical and antisymmetric fields. Despite the previous case, some of the projections are symmetrical in one direction and antisymmetric in another (the field E_x, for example, is symmetrical along $x-$direction and antisymmetric along $y-$direction, but the E_y field possesses opposite properties). The field distributions are more difficult to analyse then those of the first mode.

On the fig.5.10 the field distributions are more complex then for the 10-th mode. It is also possible to find symmetry of the field distributions (E_x is anti-symmetrical along $x-$direction and symmetrical along $y-$direction) but analysis of the fields is difficult. The field distributions are similar to the distributions of the plane waves. The same property of the distribution function for high-number modes is observed in the one-dimensional case. Thus far we can conclude that two-dimensional modes are valid.

Three-dimensional field distribution yields an estimated representation of

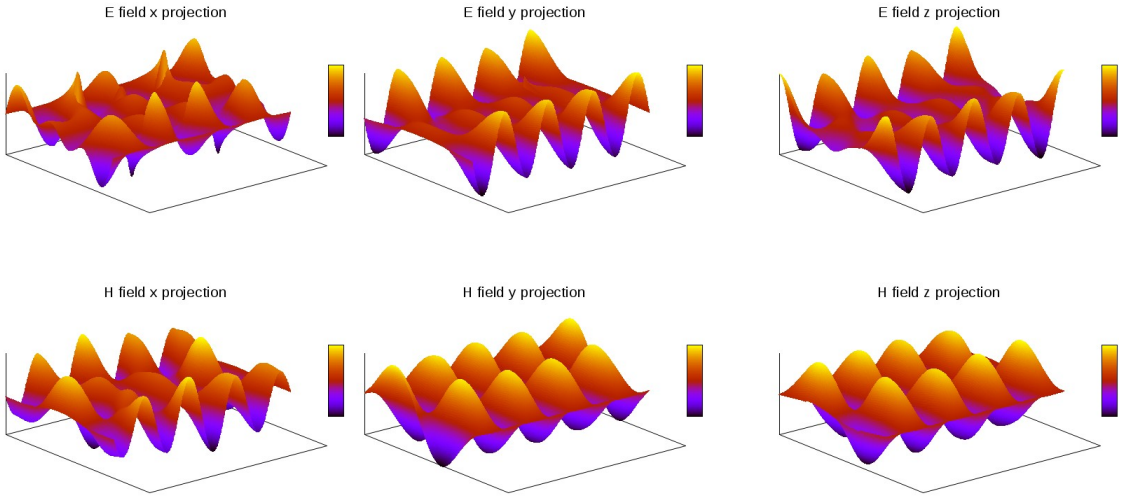


Figure 5.10: Three dimensional representation of the 101-th mode fields projections for dielectric case

the fields along the elementary cell. It allows to see the view of the disturbance at the corners or to evaluate the accuracy of the field representation for a given representation number. However, for high-order modes three dimensional images projected on the plane do not give a clear picture to characterise the field (fig.5.10). Colour coded plane projections of the field distribution are more convenient and suitable for analysis of the modal properties.

A colour coding is the technique when amplitude at certain point of the figure represents intensity or amplitude of the field at this point. Colour-encoded figures of the modal fields in dielectric case are depicted on the fig.5.11-5.13 under normal incidence.

The encoding is selected so that the black colour corresponds the the absence of the field, the red colour corresponds to the positive values and blue one to the negative.

The field intensity values are comparable within the group of field polarisations. Thus we can compare intensities of the field projections. The fields are not comparable for different figures because they are not normalised.

The fig.5.11 does not provide more information then can be obtained

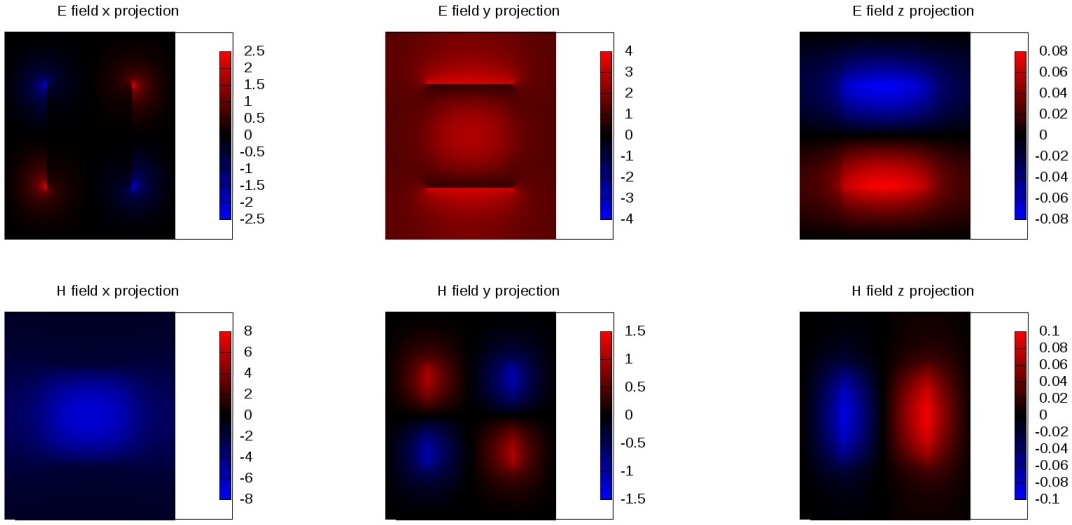


Figure 5.11: Colour encoded field distributions of the first mode

from the fig.5.8. We can definitely see that the E_x field is concentrated on the corners of the tooth outside of it. The sing of the fields is intentionally retained so we can see that H_x field is blue everywhere (negative) while E_y field is everywhere red (positive). It can be well seen that the E_y field is discontinuous along $y-$direction in the region corresponding to the teeth and continuous along $x-$direction everywhere and along $y-$direction in the region with constant refractive index. The E_y field is more intensive then E_x and E_z is approximately zero. The H_z field is also close to zero, but the H_x possesses greater amplitude.

For the 10-th mode we can see from the fig.5.12 that, unlike first mode, each field projection has comparable amplitudes. It is much easier to analyse symmetry of the field projections then in case of three-dimensional projection (fig.5.9). All the fields projections exists inside the teeth. For the E_y projection strong peaks on the corners can be observed.

Colour-encoded field representation simplifies significantly analysis of the 100-th mode. On the fig.5.13 we can see that the E_x projection is discontinuous along $x-$direction and boarders of the teeth can be seen. Analysis of the symmetry is reduced to the comparison of the colours at the boarders of

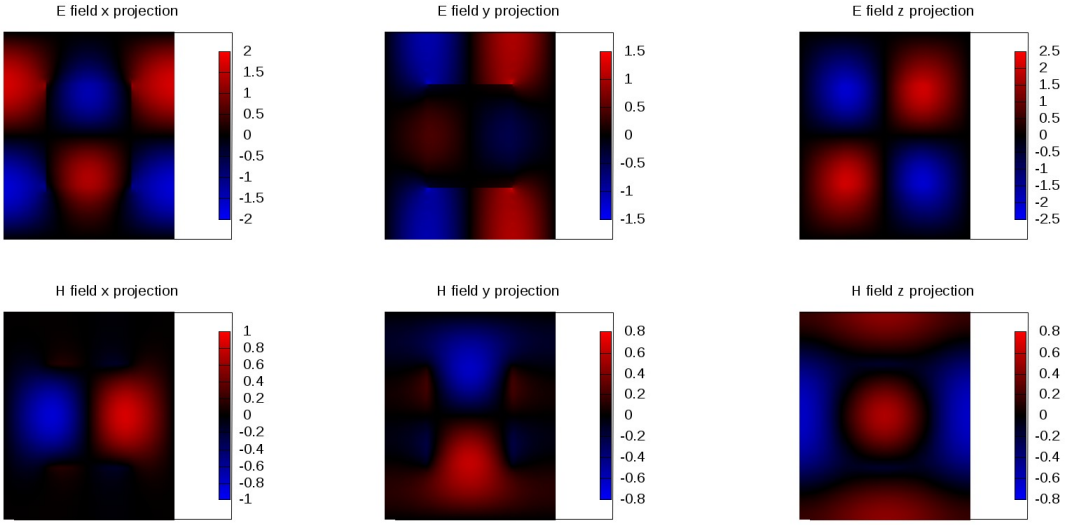


Figure 5.12: Colour encoded field distributions of the 10-th mode

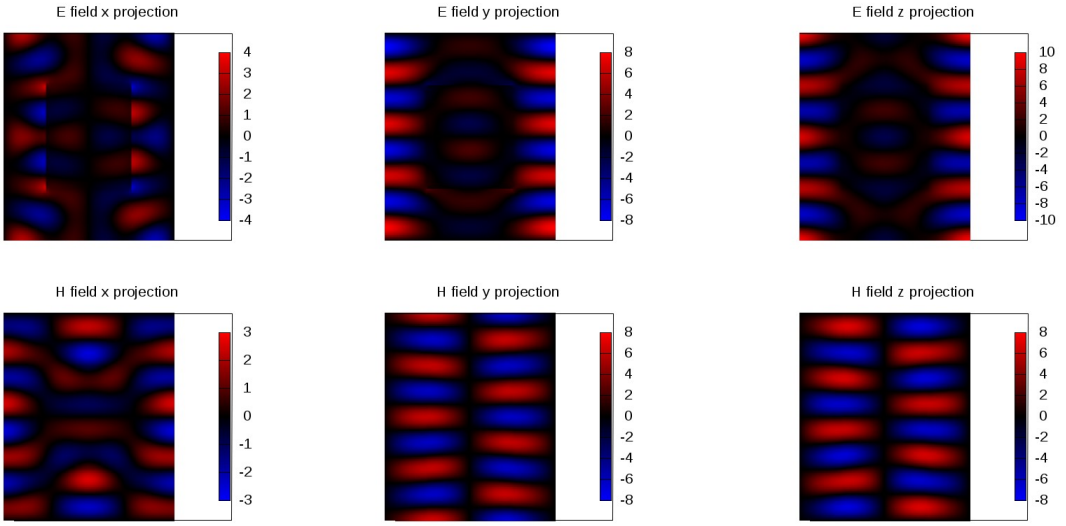


Figure 5.13: Colour encoded field distributions of the 100-th mode

each picture. For the electric field we can see that the field is concentrated mainly between the tooth long $y-$direction.

From these figures, similar to the one-dimensional case, it is possible to match between the modes and diffraction orders by comparing the parity of

the field distribution and the number of the zero intersections along each coordinate. We can conclude that the mode corresponds best to a certain diffraction order and such a correspondence is observed for all modes.

Mode composition for dielectrics in the case of 2-dimensional grating reveals an unexpected feature: there are coupled modes. Coupled modes in one-dimensional case were an exclusive property of a metal grating and the explanation of this coupling can be found at [46]. After the reconstruction of the two coupled modal fields we can not distinguish them or attribute to some diffraction order.

Taking as a linear combination of them gives a discernible and clear presentation on the properties of the coupled modes.

In contrast to the modes of one-dimensional grating, diffraction orders for the field components have the same wave vector projections for both polarisations. And, in case of normal incidence one value of the k^z projection may correspond up to the 8 combinations: twice the $x-$projection times twice $y-$direction and times two for polarisation state. From the one-dimensional grating research it is expected that to the number r of waves with k^z value will correspond required number of modes and equal to r. If we look at the density of the β^2 and the k^{z2} diffraction order projections we can conclude that the modal propagation constants are grouped around the orders projection values with a shift towards more negative values and sometimes with overlapping groups in close (but different) positions of the diffraction orders. Groups corresponding to several projection numbers with the same value are mixed and sequence matching modes and orders is not preserved. In the case of normal incidence and the strong difference in the periods over different coordinates empty area on the axis can occur. When the nearest order is far and all the modes corresponding to the previous orders are located in the area are still far from the region influenced by the next order. For the modelling and selection of the mode number to be taken into consideration I suggest on the basis of this observation to use of the number of modes which correspond to such a cutoff. It allows to obtain the best match between modes and orders and error introduces by the truncation of the matrix is minimal for a given number of modes. Based on the density of the diffraction orders

we can also evaluate the correctness of the found propagation constants in terms of completeness.

The minimum divergence of the order and modal total numbers is the criterion of the basis completeness, that all the necessary propagation vector constants of the two-dimensional diffraction problem are found.

5.2 Diffraction efficiency calculation based on the obtained field distributions and propagation constants

5.2.1 Steps required to build scattering matrix

Making the propagation constants and the corresponding field distributions we match fields on the grating surfaces. Tangential components of the magnetic and electrical field are continuous at each side of the boundary. In the introduced geometry that is expressed like:

$$\begin{aligned} E_x(x,y,t+0) &= E_x(x,y,t-0) \quad H_x(x,y,t+0) = H_x(x,y,t-0) \\ E_y(x,y,t+0) &= E_y(x,y,t-0) \quad H_y(x,y,t+0) = H_y(x,y,t-0) \end{aligned}$$

$$\begin{aligned} E_x(x,y,b-0) &= E_x(x,y,b+0) \quad H_x(x,y,b-0) = H_x(x,y,b+0) \\ E_y(x,y,b-0) &= E_y(x,y,b+0) \quad H_y(x,y,b-0) = H_y(x,y,b+0) \end{aligned}$$

where t is the coordinate of the top-surface of the two-dimensional grating, and b is the bottom surface coordinate. Exact numbers of these values depend on the position of the zero point on the $z-$axis. Independent on the definition point equation $depth = t - b$ is valid for these coordinates.

Expressing the fields with theirs modal or order representation and consequent application of the orthogonal operator leads to the set of equations where amplitudes at one side of the boundary are expressed as a sum of the linear operator of the field representation from another side of the boundary.

Analytical expressions allows to reduce equation to the form when each field amplitudes is expressed as a function of all the amplitudes from another side. These expression can be rewritten in a matrix form, building transition matrix.

The previously defined technique (2.2.3) to construct scattering matrix based on the two transition matrices, written for each of the grating boundary, is then used.

5.2.2 Field expressions

The field above the grating in the region with permittivity ε_I consists of the waves going down to the grating surface and going upwards. Field expression is similar to the one for the one-dimensional case (3.6) with the only difference that each component is now denoted with "coordinate number" containing two indices (n, m). We have already truncated the field decomposition so that there are N is the maximal order projection on the $x-$axis and M is the maximum for the $y-$axis:

$$E_x(x,y,z) = \sum_{n=-N}^{N}\sum_{m=-M}^{M} \frac{k_n^x}{\gamma_{n,m}} O_{n,m}^{+e} - \frac{1}{\omega\varepsilon_I}\frac{k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{+h} + \frac{k_n^x}{\gamma_{n,m}} O_{n,m}^{-e} + \frac{1}{\omega\varepsilon_I}\frac{k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-h}$$

$$E_y(x,y,z) = \sum_{n=-N}^{N}\sum_{m=-M}^{M} -\frac{k_m^y}{\gamma_{n,m}} O_{n,m}^{+e} - \frac{1}{\omega\varepsilon_I}\frac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{+h} - \frac{k_m^y}{\gamma_{n,m}} O_{n,m}^{-e} + \frac{1}{\omega\varepsilon_I}\frac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-h}$$

$$H_x(x,y,z) = \sum_{n=-N}^{N}\sum_{m=-M}^{M} -\frac{1}{\omega\mu}\frac{k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{+e} + \frac{k_n^x}{\gamma_{n,m}} O_{n,m}^{+h} + \frac{1}{\omega\mu}\frac{k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-e} + \frac{k_n^x}{\gamma_{n,m}} O_{n,m}^{-h}$$

$$H_y(x,y,z) = \sum_{n=-N}^{N}\sum_{m=-M}^{M} -\frac{1}{\omega\mu}\frac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{+e} - \frac{k_m^y}{\gamma_{n,m}} O_{n,m}^{+h} + \frac{1}{\omega\mu}\frac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-e} - \frac{k_m^y}{\gamma_{n,m}} O_{n,m}^{-h}$$

While superscript of the order amplitude O stands for TE-polarised order (e) and TM-polarised (h) order; the sign of the superscript denotes propagation direction along the $z-$axis. Dependence on the $x-$ and $y-$axes in the form $\exp(jk_n^x x) \cdot \exp(jk_m^y y)$ is omitted due to the equality for each term while the $z-$axis dependence is in the form $\exp(\pm jk_{n,m}^z z)$ and sign is taken the same as one in the superscript.

Expression of the field interior of the grating region depends on the method, selected for the propagation research procedure. For the simplicity of this step description we will suppose that the field component are linear operator of some form from the mode (propagation constant and correspond-

ing amplitudes vector):

$$
\begin{aligned}
E_x(x,y,z) &= \sum_{v=0}^{V} A_v^+ \mathcal{L}_x^e \left(\psi_v(x,y)\exp(j\beta_v z)\right) + A_v^- \mathcal{L}_x^e \left(\psi_v(x,y)\exp(-j\beta_v z)\right) \\
E_y(x,y,z) &= \sum_{v=0}^{V} A_v^+ \mathcal{L}_y^e \left(\psi_v(x,y)\exp(j\beta_v z)\right) + A_v^- \mathcal{L}_y^e \left(\psi_v(x,y)\exp(-j\beta_v z)\right) \\
H_x(x,y,z) &= \sum_{v=0}^{V} A_v^+ \mathcal{L}_x^h \left(\psi_v(x,y)\exp(j\beta_v z)\right) + A_v^- \mathcal{L}_x^h \left(\psi_v(x,y)\exp(-j\beta_v z)\right) \\
H_y(x,y,z) &= \sum_{v=0}^{V} A_v^+ \mathcal{L}_y^h \left(\psi_v(x,y)\exp(j\beta_v z)\right) + A_v^- \mathcal{L}_y^h \left(\psi_v(x,y)\exp(-j\beta_v z)\right)
\end{aligned}
$$

here introduced amplitudes A of the mode numbered v. Superscript of the amplitudes denotes the direction of the mode propagation. Field operators $\mathcal{L}_j^p$ returns projection j of the polarisation p component of the mode. Operator view does not depend on the modal number v but depends on the underlying method used.

5.2.3 Orthogonal operator application

Reexpressing the field equations is done by application of the consequent orthogonal operator to both sides of the equations and leads from the form where at both sides of the equation we have sum to the form where only one of the sides will have summation. An operator orthogonal to one of the basis (modal or plane wave) can be applied. For the two-dimensional case applying orthogonal to the plane wave decomposition operator is straightforward. Operator corresponding to the coordinate number n', m' has the following form:

$$
\mathcal{O}_{n',m'}(f(x,y)) = \frac{1}{\Lambda_x}\frac{1}{\Lambda_y}\int_0^{\Lambda_x}\int_0^{\Lambda_y} f(x,y)\cdot\exp(-jk_{n'}^x x)\exp(-jk_{m'}^y y)dxdy
$$

Operator defined this way also poses normalisation property:

$$
\mathcal{O}_{n',m'}\left(\exp(jk_n^x x)\exp(jk_m^y y)\right) = \delta_{n=n',m=m'}
$$

After applying operator we get the following system for each pair (n, m):

$$\begin{aligned}
&\tfrac{k_m^y}{\gamma_{n,m}} O_{n,m}^{+e} - \tfrac{1}{\omega\varepsilon_I} \tfrac{k_n^x k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{+h} + \tfrac{k_m^y}{\gamma_{n,m}} O_{n,m}^{-e} + \tfrac{1}{\omega\varepsilon_I} \tfrac{k_n^x k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-h} = \\
&\sum_{v=0}^{V} A_v^+ \mathcal{O}_{n,m} \left(\mathcal{L}_x^{+e} \left(\psi_v(x,y)\right)\right) + A_v^- \mathcal{O}_{n,m} \left(\mathcal{L}_x^{-e} \left(\psi_v(x,y)\right)\right) = E_1 \\
&-\tfrac{k_n^x}{\gamma_{n,m}} O_{n,m}^{+e} - \tfrac{1}{\omega\varepsilon_I} \tfrac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{+h} - \tfrac{k_n^x}{\gamma_{n,m}} O_{n,m}^{-e} + \tfrac{1}{\omega\varepsilon_I} \tfrac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-h} = \\
&\sum_{v=0}^{V} A_v^+ \mathcal{O}_{n,m} \left(\mathcal{L}_y^{+e} \left(\psi_v(x,y)\right)\right) + A_v^- \mathcal{O}_{n,m} \left(\mathcal{L}_y^{-e} \left(\psi_v(x,y)\right)\right) = E_2 \\
&+\tfrac{1}{\omega\mu} \tfrac{k_{n,m}^z k_n^x}{\gamma_{n,m}} O_{n,m}^{+e} + \tfrac{k_m^y}{\gamma_{n,m}} O_{n,m}^{+h} - \tfrac{1}{\omega\mu} \tfrac{k_n^x k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-e} + \tfrac{k_m^y}{\gamma_{n,m}} O_{n,m}^{-h} = \\
&\sum_{v=0}^{V} A_v^+ \mathcal{O}_{n,m} \left(\mathcal{L}_x^{+h} \left(\psi_v(x,y)\right)\right) + A_v^- \mathcal{O}_{n,m} \left(\mathcal{L}_x^{-h} \left(\psi_v(x,y)\right)\right) = E_3 \\
&+\tfrac{1}{\omega\mu} \tfrac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{+e} - \tfrac{k_n^x}{\gamma_{n,m}} O_{n,m}^{+h} - \tfrac{1}{\omega\varepsilon_I} \tfrac{k_m^y k_{n,m}^z}{\gamma_{n,m}} O_{n,m}^{-e} - \tfrac{k_n^x}{\gamma_{n,m}} O_{n,m}^{-h} = \\
&\sum_{v=0}^{V} A_v^+ \mathcal{O}_{n,m} \left(\mathcal{L}_y^{+h} \left(\psi_v(x,y)\right)\right) + A_v^- \mathcal{O}_{n,m} \left(\mathcal{L}_y^{-h} \left(\psi_v(x,y)\right)\right) = E_4
\end{aligned}$$

It should be mentioned that operators $\mathcal{L}_j^p$ and $\mathcal{O}_{n,m}$ do not commute and should be applied in the correct order. The right part of each line will be later denoted as E_i to simplify the equations. We may obtain the expression for each plane wave amplitude after some arithmetical exercise. We arrive at the expressions:

$$\begin{aligned}
O_{n,m}^{+e} &= \tfrac{1}{2\gamma_{n,m}} \left(k_m^y E_1 - k_n^x E_2 + \tfrac{\omega\mu}{k_{n,m}^z} \left(k_n^x E_3 + k_m^y E_4\right)\right) \\
O_{n,m}^{+h} &= \tfrac{1}{2\gamma_{n,m}} \left(k_m^y E_3 - k_n^x E_4 - \tfrac{\omega\varepsilon_I}{k_{n,m}^z} \left(k_n^x E_1 + k_m^y E_2\right)\right) \\
O_{n,m}^{-e} &= \tfrac{1}{2\gamma_{n,m}} \left(k_m^y E_1 - k_n^x E_2 - \tfrac{\omega\mu}{k_{n,m}^z} \left(k_n^x E_3 + k_m^y E_4\right)\right) \\
O_{n,m}^{-h} &= \tfrac{1}{2\gamma_{n,m}} \left(k_m^y E_3 - k_n^x E_4 + \tfrac{\omega\varepsilon_I}{k_{n,m}^z} \left(k_n^x E_1 + k_m^y E_2\right)\right)
\end{aligned} \tag{5.9}$$

5.2.4 Transition matrix components

All the grating specific parts are hidden in the E_i values of the (5.9). Each value in its' turn is a sum over modal amplitudes. The final expression of the order amplitudes contain linear combination of left-side sums. Plane waves

amplitudes can be re-expressed in a matrix form:

$$\begin{pmatrix} O^{+e} \\ O^{+h} \\ O^{-e} \\ O^{-h} \end{pmatrix} = T_{2d} \begin{pmatrix} A^{+} \\ A^{-} \end{pmatrix}$$

where as each amplitude a vector of amplitudes is denoted, that is $A^+ = \left(A_0^+, A_1^+ \dots A_V^+\right)^T$ and same for the others. Truncated double indexing of the order amplitudes can be reordered so that uniformly coordinate number can be obtained from the index, or index corresponding to the coordinate number.

The $T-$matrix components corresponding to the notation above consist of the following sub-matrices:

$$T = \begin{bmatrix} T_+^{+e} & T_-^{+e} \\ T_+^{+h} & T_-^{+h} \\ T_+^{-e} & T_-^{-e} \\ T_+^{-e} & T_-^{-e} \end{bmatrix}$$

With an index j uniformly mapped to the coordinate number (n, m) and containing information about both n and m, the T matrix components are:

$$\begin{aligned} \left[T_\pm^{+e}\right]_{ij} = & \quad \tfrac{1}{2\gamma_j}\Big[k^y_{m(j)}\mathcal{O}_j\left(\mathcal{L}_x^{\pm e}\left(\psi_i(x,y)\right)\right) - k^x_{n(j)}\mathcal{O}_j\left(\mathcal{L}_y^{\pm e}\left(\psi_i(x,y)\right)\right) \\ & + \tfrac{\omega\mu}{k_j^z}\left(k^x_{n(j)}\mathcal{O}_j\left(\mathcal{L}_x^{\pm h}\left(\psi_i(x,y)\right)\right) + k^y_{m(j)}\mathcal{O}_j\left(\mathcal{L}_y^{\pm h}\left(\psi_i(x,y)\right)\right)\right)\Big] \end{aligned} \tag{5.10}$$

$$\begin{aligned} \left[T_\pm^{-e}\right]_{ij} = & \quad \tfrac{1}{2\gamma_j}\Big[k^y_{m(j)}\mathcal{O}_j\left(\mathcal{L}_x^{\pm e}\left(\psi_i(x,y)\right)\right) - k^x_{n(j)}\mathcal{O}_j\left(\mathcal{L}_y^{\pm e}\left(\psi_i(x,y)\right)\right) \\ & - \tfrac{\omega\mu}{k_j^z}\left(k^x_{n(j)}\mathcal{O}_j\left(\mathcal{L}_x^{\pm h}\left(\psi_i(x,y)\right)\right) + k^y_{m(j)}\mathcal{O}_j\left(\mathcal{L}_y^{\pm h}\left(\psi_i(x,y)\right)\right)\right)\Big] \end{aligned} \tag{5.11}$$

$$\begin{aligned} \left[T_\pm^{+h}\right]_{ij} = & \quad \tfrac{1}{2\gamma_j}\Big[k^y_{m(j)}\mathcal{O}_j\left(\mathcal{L}_x^{\pm h}\left(\psi_i(x,y)\right)\right) - k^x_{n(j)}\mathcal{O}_j\left(\mathcal{L}_y^{\pm h}\left(\psi_i(x,y)\right)\right) \\ & - \tfrac{\omega\varepsilon_I}{k_j^z}\left(k^x_{n(j)}\mathcal{O}_j\left(\mathcal{L}_x^{\pm e}\left(\psi_i(x,y)\right)\right) + k^y_{m(j)}\mathcal{O}_j\left(\mathcal{L}_y^{\pm e}\left(\psi_i(x,y)\right)\right)\right)\Big] \end{aligned} \tag{5.12}$$

$$\left[T_{\pm}^{-h}\right]_{ij} = \frac{1}{2\gamma_j}\left[k_{m(j)}^{y}\mathcal{O}_j\left(\mathcal{L}_x^{\pm h}\left(\psi_i(x,y)\right)\right) - k_{n(j)}^{x}\mathcal{O}_j\left(\mathcal{L}_y^{\pm h}\left(\psi_i(x,y)\right)\right) + \frac{\omega\varepsilon_I}{k_j^z}\left(k_{n(j)}^{x}\mathcal{O}_j\left(\mathcal{L}_x^{\pm e}\left(\psi_i(x,y)\right)\right) + k_{m(j)}^{y}\mathcal{O}_j\left(\mathcal{L}_y^{\pm e}\left(\psi_i(x,y)\right)\right)\right)\right] \tag{5.13}$$

where sign in the under-script of the T stands for the modal amplitude direction (going up is a plus and going down is a minus). j is the row of the matrix and i is the column. Thus far matrix operation is valid.

5.2.5 Scattering matrix

The transition matrix for the second boundary of the two-dimensional grating is constructed according to the same formulae with the only difference that ε_I should be replaced with the permittivity of the semi-infinite media at the second boundary ε_{II}. After completing of this operation we have two matrices with height equal $2 \cdot N_x \cdot M_y$ (where N_x is the total number of the $x-$axis order projections taken and M_y is the total number of the $y-$ projections of the orders) and width equal $2 \cdot V$ where V is total number of the modes used.

To use scattering matrix procedure described in the sec.2.2.2 we need to get an inverse matrix. The operation of inversion is defined only for the square matrices. From this point we come to the fact that $2N_x \cdot M_y$ should be equal to V so that the procedure can be applied. We are dealing with the modal method so the modal amplitude at an arbitrary depth inside the grating can be calculated by multiplication by the phase term $\exp(j\beta d)$ where d is counted starting from the point of modal amplitude definition.

Scattering matrix build on the transition matrices defined above will set the correspondence between the $j-$th order to the $k-$th order. For both values j and k inverse function to obtain coordinate number are applied and backwards. To get some information regarding the order with the n, m coordinate number we should find corresponding to that value number j and k.

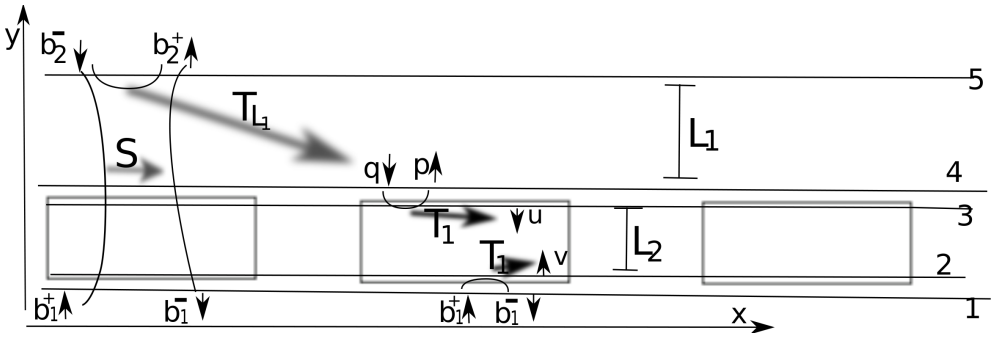


Figure 5.14: Figure with layer description and comments on the amplitudes definition

5.2.6 Circle-locked orders

If we introduce correspondence function which has inverse function we can shift from the square-based definition of the orders (when for each projection on the $x-$ there is a complete set of the $y-$projections) to the circle-locked definition. We will take into consideration only the orders so that the k_j^z projections are not greater then some value β_{max} or as close as possible to that value. That complexity brings benefits because we do not take orders with high coordinate numbers (with the biggest k^z values) when there is no corresponding modes to them. As it was shown in the [46], accuracy does not increase with the number of modes taken into account if there is one or several orders without corresponding modes.

5.2.7 Example of the method application

Here we will consider already introduced structure. Layer with homogeneous permittivity is the region between lines 4 and 5 also below line 1. Regions between lines 1-2 and 3-4 are just boundaries with zero depth. And one-dimensional grating is inside the region limited with lines 2 and 3 From the previous step we a set of the propagation constants $\{\pm\beta_i\}$ and eigenvectors of the dispersion matrix $\left\{\vec{P^i}\right\}$corresponding to them. The vector $\vec{P^i}$ contains amplitudes of the orders at the boundaries of an elementary stripe. They are at the lines 1 going upwards (b_1^+) and at line 5 going downwards (b_2^-) of the fig.5.2.7. We can obtain vector amplitudes of the b_2^+ and b_1^- with a

multiplication of the layer scattering matrix by eigenvector.

$$\vec{P^i} = \begin{pmatrix} \vec{b_1^+} \\ \vec{b_2^-} \end{pmatrix}$$

$$\begin{pmatrix} \vec{b_1^-} \\ \vec{b_2^+} \end{pmatrix} = S \times \begin{pmatrix} \vec{b_1^+} \\ \vec{b_2^-} \end{pmatrix}$$

From the transition matrix of the layer L_1 we can get the amplitudes on the other boundary of the layer L_1. We denote amplitudes going downwards as q and upwards as p.

$$\begin{pmatrix} \vec{p} \\ \vec{q} \end{pmatrix} = T_{L_1} \times \begin{pmatrix} \vec{b_2^+} \\ \vec{b_2^-} \end{pmatrix}$$

For both of the one-dimensional grating boundaries we have transition from the space to grating matrix. We can find thus far the modal amplitudes sets u and v according to the

$$\begin{pmatrix} \cdot \\ \vec{u} \end{pmatrix} = T_1 \begin{pmatrix} \vec{p} \\ \vec{q} \end{pmatrix}$$
$$\begin{pmatrix} \vec{v} \\ \cdot \end{pmatrix} = T_1 \begin{pmatrix} \vec{b_1^+} \\ \vec{b_1^-} \end{pmatrix}$$

This is the procedure used to restore field amplitudes of a particular two-dimensional mode i. The field over the elementary layer are defined differently. As a consequence the field operators are different. Each amplitudes set can be split into the subsets of the TE- and TM-amplitudes. From the assumption that we used R as a representation number (there is R components of each polarisation in each set) we may write the field operators necessary for the scattering matrix calculation ($\mathcal{L}_x^{\pm e}$, $\mathcal{L}_x^{\pm h}$, $\mathcal{L}_y^{\pm e}$, $\mathcal{L}_y^{\pm h}$):

$$\mathcal{L}_x^{\pm e}\left(\psi_i(x,y)\right) =$$
$$\begin{cases} \sum\limits_{r=-R/2}^{R/2} \left(\mp \frac{\beta_i}{\gamma_r} p_r^e(y) - \frac{1}{\omega\varepsilon_2} \frac{k_r^y k_r^x}{\gamma_r} p_r^h(y) \mp \frac{\beta_i}{\gamma_r} q_r^e(y) + \frac{1}{\omega\varepsilon_2} \frac{k_r^x k_r^y}{\gamma_r} q_r^h(y) \right) f_r(x), & mr_y\Lambda_y < y \leq \Lambda_y \\ \sum\limits_{r=0}^{R} \frac{\zeta_r}{\omega\varepsilon_2} \phi_r^h(x) v_r^h(y) + \frac{\zeta_r}{\omega\varepsilon_2} \phi_r^h(x) u_r^h(y), & 0 < y \leq mr_y\Lambda_y \end{cases}$$

$$\mathcal{L}_y^{\pm e}\left(\psi_i(x,y)\right) =$$
$$\begin{cases} \sum_{r=-R/2}^{R/2} \left(-\frac{1}{\omega\varepsilon_2}\gamma_r p_r^h(y) - \frac{1}{\omega\varepsilon_2}\gamma_r q_r^h(y)\right) f_r(x), & mr_y\Lambda_y < y \leq \Lambda_y \\ \sum_{r=0}^{R} \pm\frac{\beta_i}{\zeta_r}\phi_r^e(x)v_r^e(y) - j\frac{k_r^y}{\zeta_r\omega\varepsilon_2}\phi_r^{h\prime}(x)v_r^h(y) \pm \frac{\beta_i}{\zeta_r}\phi_r^e(x)u_r^e(y) + j\frac{k_r^y}{\zeta_r\omega\varepsilon_2}\phi_r^{h\prime}(x)u_r^h(y), & \\ & 0 < y \leq mr_y\Lambda_y \end{cases}$$

$$\mathcal{L}_x^{\pm h}\left(\psi_i(x,y)\right) =$$
$$\begin{cases} \sum_{r=-R/2}^{R/2} \left(\frac{k_r^y k_r^x}{\omega\mu_0\gamma_r}p_r^e(y) \mp \frac{\beta_i}{\gamma_r}p_r^h(y) - \frac{k_r^y k_r^x}{\omega\mu_0\gamma_r\gamma_r}q_r^e(y) \mp \frac{\beta_i}{\gamma_r}q_r^h(y)\right) f_r(x), & mr_y\Lambda_y < y \leq \Lambda_y \\ \sum_{r=0}^{R} -\frac{\zeta_r}{\omega\mu_0}\phi_r^e(x)v_r^e(y) + -\frac{\zeta_r}{\omega\mu_0}\phi_r^e(x)u_r^e(y), & 0 < y \leq mr_y\Lambda_y \end{cases}$$

$$\mathcal{L}_y^{\pm h}\left(\psi_i(x,y)\right) =$$
$$\begin{cases} \sum_{r=-R/2}^{R/2} \left(\frac{\gamma_r}{\omega\mu_0}p_r^e(y) + \frac{\gamma_r}{\omega\mu_0}q_r^e(y)\right) f_r(x) & mr_y\Lambda_y < y \leq \Lambda_y \\ \sum_{r=0}^{R} j\frac{k_r^y}{\zeta_r\omega\mu_0}\phi_r^{e\prime}(x)v_r^e(y) \pm \frac{\beta_i}{\zeta_r}\phi_r^h(x)v_r^h(y) - j\frac{k_r^y}{\zeta_r\omega\mu_0}\phi_r^{e\prime}(x)u_r^e(y) \pm \frac{\beta_i}{\zeta_r}\phi_r^h(x)u_r^h(y) & \\ & 0 < y \leq mr_y\Lambda_y \end{cases}$$

where $\gamma_r = \sqrt{\beta^2 + {k_r^x}^2}$, $k_r^y = \sqrt{\varepsilon_2\mu_0\omega^2 - {\gamma_r}^2}$, $f_r(x) = \exp(jk_r^x x)$, $p_r(y) = p_r\exp(jk_r^y y)$, $q_r(y) = q_r\exp(-jk_r^y y)$, $v_r^e(y) = v_r^e\exp(jk_r^e y)$, $u_r^e(y) = u_r^e\exp(-jk_r^e y)$, $v_r^h(y) = v_r^h\exp(jk_r^h y)$, $u_r^h(y) = u_r^h\exp(-jk_r^h y)$ and $k_r^p = \sqrt{\zeta_r^{p2} - \beta^2}$, where ζ_r^p is the propagation constant of the mode r with polarisation p with distribution function $\phi_r^p(x)$. The result of the orthogonal operator to each of these functions will be:

$$\mathcal{O}_j\left(\mathcal{L}(\psi_i(x,y))\right) = \frac{1}{\Lambda_x}\frac{1}{\Lambda_y}\int_{y=0}^{\Lambda_y}\int_{x=0}^{\Lambda_x}\exp(-jk_{m(j)}^y y)\exp(-jk_{n(j)}^x x)\mathcal{L}(\psi_i(x,y))dydx$$

taking into account that function $\psi(x,y) = g(x)\cdot f(y)$ is separable along

coordinate dependence we can separate integration over the axes:

$$\mathcal{O}_j\left(\mathcal{L}(\psi_i(x,y))\right) = \frac{1}{\Lambda_x}\frac{1}{\Lambda_y}\left\{\begin{array}{c} \int\limits_{y=0}^{mr_y\Lambda_y}\int\limits_{x=0}^{\Lambda_x} \exp(-jk^y_{m(j)}y)\exp(-jk^x_{n(j)}x)\mathcal{L}(\psi_i(x,y))dydx \\ + \int\limits_{y=mr_y\Lambda_y}^{\Lambda_y}\int\limits_{x=0}^{\Lambda_x} \exp(-jk^y_{m(j)}y)\exp(-jk^x_{n(j)}x)\mathcal{L}(\psi_i(x,y))dydx \end{array}\right.$$
$$= \mathcal{O}^i_j(L_2) + \mathcal{O}^i_j(L_1)$$

First we will consider operator result on the homogeneous region. Field is represented there with the plane wave superposition. Integral over the $x-$axis from the sum becomes only the single term:

$$\mathcal{O}^i_j(L_1) = \frac{1}{\Lambda_x}\frac{1}{\Lambda_y}\int\limits_{y=mr_y\Lambda_y}^{\Lambda_y}\int\limits_{x=0}^{\Lambda_x} \exp(-jk^y_{m(j)}y)\exp(-jk^x_{n(j)}x)\times$$
$$\sum_{r=-R/2}^{R/2}\left(C_1p^e_r(y) + C_2p^h_r(y) + C_3q^e_r(y) + C_4q^h_r(y)\right)\exp(jk^x_rx)dydx =$$
$$\delta_{n(j)=r}\left(\frac{1}{\Lambda_y}\int\limits_{y=mr_y\Lambda_y}^{\Lambda_y} C_1p^e_r(y) + C_2p^h_r(y) + C_3q^e_r(y) + C_4q^h_r(y)dy\right) =$$
$$\delta_{n(j)=r}\left(C_1I^{++}_{i,j}p^e_r + C_2I^{++}_{i,j}p^h_r + C_3I^{+-}_{i,j}q^e_r + C_4I^{+-}_{i,j}q^h_r\right)$$

where the following integrals are used:

$$I^{++}_{i,j} = \frac{1}{\Lambda_y}\int\limits_{y=mr_y\Lambda_y}^{\Lambda_y} \exp(-jk^y_{m(j)}y)\exp(jk^y_{i,j}y)dy$$
$$I^{+-}_{i,j} = \frac{1}{\Lambda_y}\int\limits_{y=mr_y\Lambda_y}^{\Lambda_y} \exp(-jk^y_{m(j)}y)\exp(-jk^y_{i,j}y)dy$$

where $k^y_{n(j)} = \sqrt{\omega^2 - \left({k^x_{n(j)}}^2 + \beta^2_i\right)}$, $k^y_{i,j}$ is the $y-$projection of the order r for which the equation $m(j) = r$ is satisfied. Thus far notation (i,j) of the $i-$ mode for the $j-$th order is used. Coefficients C_s depend on the field polarisation and projection, which are expressed, and represented with combination of the wave vector projections of the order inside the grating. The expressions for $\mathcal{O}^i_j(L_1)$ can be found analytically reducing calculation difficulty of the problem.

There is no trivial way to get result of the operator action for the region

L_2. We need to calculate all the integrals but taken into account for the separable functions and linearity of the operator we can write:

$$\int_{x=0}^{\Lambda_x} \int_{y=0}^{mr_y\Lambda_x} \sum_{r=0}^{R} f_r(x)g_r(y) = \sum_{r=0}^{R} \int_{x=0}^{\Lambda_x} f_r(x)dx \int_{y=0}^{mr_y\Lambda_x} g_r(y)dy$$

We can rewrite the operator expression in the form:

$$\mathcal{O}_j^i(L_1) = \sum_{r=0}^{R} \left(a_r^{+e}Y_r^{+e} + a_r^{-e}Y_r^{-e}\right) X_r^e + \left(a_r^{+e}Y_r^{+h} + a_r^{-e}Y_r^{-h}\right) X_r^h$$

where the integrals are of the following form:

$$X_r^p = \frac{1}{\Lambda_x} \int_{x=0}^{\Lambda_x} \mathcal{M}_{1D}\left(\psi_{1D,r}^p(x)\right) \exp(-jk_{n(j)}^x)dx$$
$$Y_r^{\pm p} = \frac{1}{\Lambda_y} \int_{y=0}^{mr_y\Lambda_y} \exp(\pm jk_{r,p}^y y) \exp(-jk_{m(j)}^y)dy$$

where expression of the X_r^p are similar to 3.8 and the choice depends on the expression (polarisation and projection of the mode), which are derived. the following variables are used: $k_{r,p}^y = \sqrt{\zeta_r^{p\,2} - \beta_i^2}$, ζ_r^p is the propagation value constant numbered r of the one dimensional problem for polarisation p.

Substituting the resulting values of the $\mathcal{O}(\mathcal{L})$ into the equations (5.10-5.13) we get transition matrices. Applying technique described in (2.2.3) we get scattering matrix of the problem.

The modal method allows to scan over the depth of the structure because the same transition matrices can be used and only propagating matrix inside the grating are different. For such scanning the structure identical to the described above is taken ($\Lambda_x = 300nm$, $\Lambda_y = 400nm$, $n_2 = 1$, $n_1 = 2.2$, $mr_x = mr_y = 0.5$), incidence is taken at $\Theta = 30$, $\phi = 40$ coming from the substrate with $n_s = 1.46$ and covering media has $n_c = 1$. All the figures represent the diffraction efficiencies taken from the incident from the substrate zeroth order and going into the zeroth reflected or transmitted orders. The polarization of the incident light for the figure 5.15 is TE, and for the 5.16

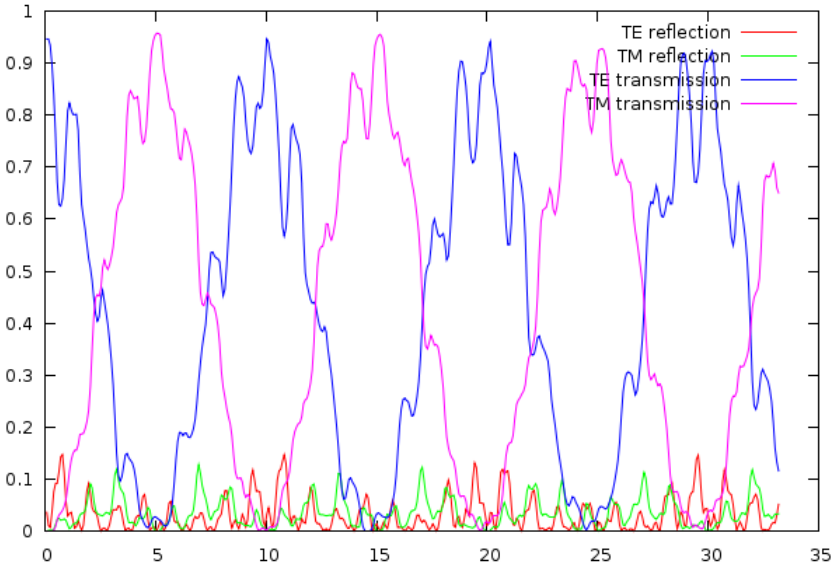


Figure 5.15: Diffraction efficiencies of the different polarizations produced by the TE-polarized incident wave

is TM. Both of these polarizarion produce TE and TM transmitted orders depending on the grating depth. The dependence is shown for the depth normalized to the wavelength.

Strange gaps on the transmission efficiencies curves can be attributed to the exitation on the edges. They are not introduced by the numerical instabilities and reproduced with the RCWA method.

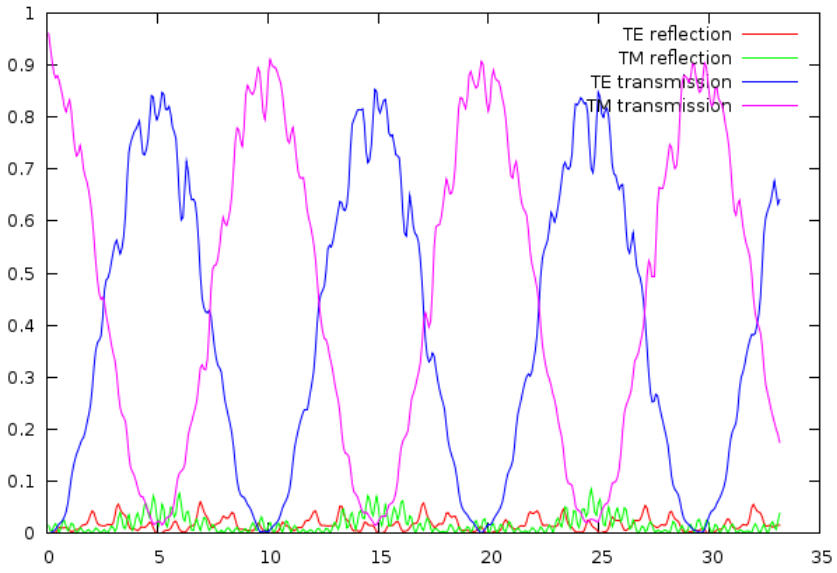


Figure 5.16: Diffraction efficiencies of the different polarizations produced by the TM-polarized incident wave

Chapter 6

Conclusion

This work can be interesting for people who starts studing of the modal methods. Already developed one-domentional methods are described in the first part of the work providing basis for further two-dimentional method development.

The developed stage of the two-dimentional modal method gives new data for analysis and already can be used for near-field investigation. As a result of this work benchmark of already existing RCWA methods was done and the true modal method showed itself as a good reference method to check and estimate the accuracy of the obtained propagation constants. Method allows to find propagation values of metal structures including metal grids with dielectric cores. The obtained propagation constants for the metal grid case show cut-off values depending on the cell sizes (the size of internal dielectric core).

The scan over depth show Fabry – Pérot interferometer's like behavior, which was already seen in case of one-dimentional gratings.

The difficulty of the method implementation can not be avoided anyway. But this difficulty is defined with the strick problem formulation and provides abilities to get as accurate results as required. The complexity of the root search procedure can be enhanced and boosted. The SIMD techniques and parallel-programming techniques can be used to reduce computation time during root-search procedure. These things can be interesting for those who

developes commercial products, because true modal method allows to calculate metal and dielectric structures without any significant influence on the algorithm. No special tricks are required to shift from the dielectrical to the metal cases.

This method promises new results and gives new approach for the metal two dimentional structures calculation. New results for the dielectrical structures can be obtained with other methods, but true modal method allows to compare results and to obtain results with demanded precision.

Bibliography

[1] G. R. J. Chandezon, D. Maystre, "A new theoretical method for diffraction gratings and its numerical application," *Jounar of Optics Paris*, vol. 11, pp. 235–241, 1980.

[2] L. R. J. W. Strutt), "On the dynamical theory of gratings," *Proc. R. Soc. London A*, vol. 79, pp. 399–416, 1907.

[3] K. Knop, "Rigorous diffraction theory for transmission phase gratings with deep rectangular grooves," *Journal of Optical Society of America*, no. 68, pp. 1206–1210, 1978.

[4] V. Karagodsky, M. C. Huang, and C. J. Chang-Hasnain, "Analytical solution and design guideline for highly reflective subwavelength gratings," *Journal of Optical Society of America*, 2008.

[5] E. Bonnet, X. Letartre, A. Cachard, A. Tishchenko, and O. Parriaux, "High resonant reflection of a confined free space beam by a high contrast segmented waveguide," *Optical and Quantum Electronics*, vol. 35, pp. 1025–1036, May 2003.

[6] R. Magnusson and M. Shokoon-Saremi, "Physical basis for wideband resonant reflectors," *Optics Express*, vol. 16, pp. 3456–3462, February 2008.

[7] C. F. Mateus, M. C. Huang, Y. Deng, A. R. Neureuther, and C. J. Chang-Hasnain, "Ultrabroadband mirror using low-index cladded subwavelength grating," *IEEE Photonics Technology Letters*, vol. 16, pp. 518–520, Febrary 2004.

[8] J. C. Pizolaro and L. G. Neto, "Phase-only optical encryption based on the zeroth-order phase-contrast technique," *Optical Engineering*, vol. 48, September 2009.

[9] D. Delbeke, R. Baets, , and P. Muys, "Polarization-selective beam splitter based on a highly efficient simple binary diffraction grating," *Applied Optics*, vol. 43, pp. 6157–6165, November 2004.

[10] D. Delbeke, P. Bienstaman, R. Bockstaele, and R. Baets, "Rigorous electromagnetic analysis of dipole emission in periodically corrugated layers: the grating-assisted resonant-cavity light-emitting diode," *Journal of Optical Society of America*, vol. 19, pp. 871–880, May 2002.

[11] D. Mawet, P. Riaud, J. Surdej, and J. Baudrand, "Subwavelength surface-relief gratings for stellar coronagraphy," *Applied Optics*, vol. 44, pp. 7313–7321, December 2005.

[12] G. Veldhuis, O. Parriaux, H. Hoekstra, and P.V.Lambeck, "Sensitivity enhancement in evanescent optical waveguide sensors," *Journal of Lightwave Thechnology*, vol. 18, pp. 677–682, May 2000.

[13] D. Dai and S. He, "Accurate two-dimensional model of an arrayed-waveguide grating demultiplexer and optimal design based on the reciprocity theory," *Journal of Optical Society of America A*, vol. 21, pp. 2392–2398, December 2004.

[14] K. Catchpole and M. Green, "A conceptual model of light coupling by pillar diffraction gratings," *Journal of Applied Physics*, vol. 101, no. 063105, 2007.

[15] A. David, H. Benisty, and C. Weisbuch, "Fast factorization rule and plane-wave expansion method for two-dimentxional photonic crystals with arbitrary hole-shape," *Physical Review B*, vol. 73, Febrary 2006.

[16] M. Docter, P. van den Berg, and I. Y. Y. Garini, "Small influence of surface plasmons on the near-field transmission intensity spatial distribution," *OSA/CLEO*, 2006.

[17] E. Popov and M. Nevière, "Grating theory: new equations in fourier space leading to fast converging results for tm polarisation," *Journal of Optical Society of America A*, vol. 17, pp. 1773–1784, 2000.

[18] M.Foresti and A. Tishchenko, "Exact solution to diffraction on a grating with parallelogramic grooves by the true modal method," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6182 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, May 2006.

[19] Wikipedia, "Maxwell's equations - wikipedia, the free encyclopedia." http://en.wikipedia.org/wiki/Maxwell%27s_equations.

[20] L. Li, "Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings," *Journal of Optical Society of America A*, vol. 13, pp. 1024–1035, May 1996.

[21] L. Li, "Justification of matrix truncation in the modal methods of diffraction gratings," *Journal of Optics: Pure Applied Optics*, vol. 1, pp. 531–536, 1999.

[22] D. Ko and J. Sambles, "Scattering matrix method for propagation of radiation in stratified media: attenuated total reflection studies of liquid crystals," *Journal of Optical Society of America A*, vol. 5, pp. 1863–1866, November 1988.

[23] A. Tishchenko, "Fenomenological representation of deep and high contrast lamellar gratings by means of the modal method," *Optical and Quantum Electronics*, vol. 37, pp. 309–330, 2005.

[24] E. Popov, M. Nevière, B. Gralak, and G. Tayeb, "Staircase approximation validity for arbitrary-shaped gratings," *Journal of Optical Society of America*, vol. 19, pp. 33–42, January 2002.

[25] R. P. (editor), M. Cadilhac, D. Maystre, P. Vincent, M. Nevière, R. McPhedran, G. Derric, and L. Botten, *Electromagnetic Theory of Gratings*. Springer-Verlag, 1980.

[26] G. Software, ""the gnu multiple precision arithmetic library"." "http://gmplib.org/".

[27] D. Goldberg, "What every computer scientist should know about floating-point arithmetic," *ACM Comput. Surv.*, vol. 23, pp. 5–48, March 1991.

[28] Wikipedia, "Fourier transform - wikipedia, the free encyclopedia." http://en.wikipedia.org/wiki/Fourier_transform.

[29] Wikipedia, "Fast fourier transform - wikipedia, the free encyclopedia." http://en.wikipedia.org/wiki/Fast_Fourier_transform.

[30] J. Bischof and K. Hehl, "Perturbation approach applied to modal diffraction methods," *Journal of Optical Society of America A*, vol. 28, pp. 859–867, May 2011.

[31] J. Chandezon, M. Dupuis, and G. Cornet, "Multicoated gratings: a differential formalism applicable in the entire optical region," *Journal of Optical Society of America*, vol. 72, pp. 839–846, July 1982.

[32] Wikipedia, "Finite-difference time-domain method - wikipedia, the free encyclopedia." http://en.wikipedia.org/wiki/Finite-difference_time-domain_method.

[33] K. Umashankar and A. Taflove, "A novel method to analyze electromagnetic scattering of complex objects," *IEEE Transactions on Electromagnetic Compatibility*, no. 24, p. 397–405, 1982.

[34] K. Yee, "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media," *IEEE Transactions on Antennas and Propagation*, no. 14, p. 302–307, 1966.

[35] A. Taflove, "Application of the finite-difference time-domain method to sinusoidal steady state electromagnetic penetration problems," *IEEE Transactions on Electromagnetic Compatibility*, no. 22, p. 191–202, 1980.

[36] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method, 3rd ed.*. Artech House Publishers, 2005.

[37] P. van den Berg and J. Fokkema, "The rayleigh hypotesis in the theory of reflection by a grating," *Journal of Optical Society of America*, vol. 69, pp. 27–31, 1979.

[38] P. van den Berg and J. Fokkema, "The rayleigh hypothesis in the theory of diffraction by a perturbation in a plane surface," *Radio Science*, vol. 15, pp. 723–732, August 1980.

[39] A. V. Tishchenko, "Numerical demonstration of the validity of the rayleigh hypothesis," *Opt. Express*, vol. 17, pp. 17102–17117, Sep 2009.

[40] A. V. Tishchenko, "Rayleigh was right: Electromagnetic fields and corrugated interfaces," *Opt. Photon. News*, vol. 21, pp. 50–54, July 2010.

[41] M. Moharam and T. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *Journal of Optical Society of America*, vol. 71, pp. 811–818, July 1981.

[42] P. Lalanne and G. M. Morris, "Highly improved convergence of the coupled-wave method for tm polarisation," *Journal of Optical Society of America A*, vol. 13, pp. 779–784, 1996.

[43] L. Li, "Use of fourier series in the analysis of discontinuous periodic structures," *Journal of Optical Society of America A*, vol. 13, pp. 1870–1876, 1996.

[44] I. Gushchin and A. V. Tishchenko, "Fourier modal method for relief gratings with oblique boundary conditions," *J. Opt. Soc. Am. A*, vol. 27, pp. 1575–1583, Jul 2010.

[45] I. Gushchin, A. V. Tishchenko, and O. Parriaux, "Light diffraction on metal gratings under conical incidence by the true-mode method," *Journal of Physics: Conference Series*, no. 139, 2008.

[46] M. Foresti, L. Menez, and A. V. Tishchenko, "Modal method in deep metal–dielectric gratings: the decisive role of hidden modes," *Journal of Optical Society of America A*, vol. 23, pp. 2501–2509, October 2006.

[47] B. Burckhardt, "Diffraction of a plane wave at a sinusoidally stratified dielectric grating," *Journal of Optical Society of America*, no. 56, pp. 1502–1509, 1966.

[48] M. G. Moharam and T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings," *Journal of Optical Society of America*, no. 72, pp. 1385–1392, 1982.

[49] S. Peng and G. Moris, "Efficient implementation of rigorous coupled-wave analysis for surface-relief gratings," *Journal of Optical Society of America*, vol. 12, pp. 1087–1096, May 1995.

[50] E. Popov and M. Nevière, "Maxwell equations in fourier space: fast-converging formulation for diffraction by arbitrary shaped, periodic, anisitropic media," *Journal of Optical Society of America A*, vol. 18, 2001.

[51] N. P. K. Cotter, T. V. Preist, and J. R. Sambles, "Scattering-matrix approach to multilayer diffraction," *Journal of Optical Society of America A*, no. 12, pp. 1097–1103, 1995.

[52] T. Schuster, J. Rouff, N. Kerwien, S. Rafler, and W. Osten, "Normal vector method for convergence improvement using the rcwa for crossed gratings," *Journal of Optical Society of America A*, vol. 24, pp. 2880–2890, September 2007.

[53] T. Weiss, G. Granet, N. A.Gippius, S. G. Tikhodeev, and H. Giessen, "Matched coordinates and adaptive spatial resolution in the fourier modal method," *Optics Express*, vol. 17, pp. 8051–8061, May 2009.

[54] G. Granet and B. Guizal, "Efficient implementation of the coupled-wave method for metallic lamellar gratings in tm polarization," *Journal of Optical Society of America A*, vol. 13, pp. 1019–1023, 1996.

[55] G. Granet, "Reformulation of the lamellar grating problem through the concept of adaptive spatial resolution," *Journal of Optical Society of America A*, vol. 16, pp. 2510–2516, 1999.

[56] B. Guizal, H. Yala, and D. Felbacq, "Reformulation of the eigenvalue problem in the fourier modal method with spatial adaptive resolution," *Optic Letters*, vol. 34, pp. 2790–2792, 2009.

[57] N. M. Lyndin, O. Parriaux, and A. V. Tishchenko, "Modal analysis and suppression of the fmm instabilities in highly conductive gratings," *Journal of Optical Society of America A*, vol. 24, pp. 3781–3788, 2007.

[58] M. Moharam, E. Grann, D. Pommet, and T. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings," *Journal of Optical Society of America A*, vol. 5, no. 12, pp. 1068–1076, 1995.

[59] P. Lalanne, "Improved formulation of the coupled-wave method for two-dimensional gratings," *Journal of Optical Society of America A*, vol. 14, pp. 1592–1598, July 1997.

[60] L. Li, "New formulation of the fourier modal method for crossed surface-relief gratings," *Journal of Optical Society of America A*, vol. 14, no. 10, pp. 2758–2767, 1997.