



UNIVERSITA' di TRENTO

Dipartimento di Fisica
Ph.D. degree in Physics

**Modeling the interaction of light with photonic structures by
direct numerical solution of Maxwell's equations**

Candidate:
Alessandro Vaccari

Advisor:
Dott. Lucia Calliari
Fondazione Bruno Kessler
FBK-LISC

*A Maria Angela, mia mamma,
che c'era quando tutto questo è iniziato ...
e c'è ancora, ma come
il sole in una mattinata di primavera!*

Contents

Introduction and motivation	vii
I Some selected topics in electromagnetism and wave optics	1
1 Electromagnetism in matter	3
1.1 Maxwell's equations in matter	3
1.2 Temporally dispersive media.	6
1.3 Solution for the sphere in the field of a plane wave	7
2 Spatial Fourier analysis and far-fields	15
2.1 Plane wave spectrum representation	15
2.2 Far-field in the angular spectrum representation	17
2.3 Kirchoff formula and the near to far field transform	18
3 Gaussian beams	23
3.1 Paraxial approximation	23
3.2 Gaussian laser beams	24
3.3 Focused beams	25
4 Planarly multilayered media	31
4.1 Matrix method for \mathcal{R} and \mathcal{T}	31
4.2 \vec{E} and \vec{H} field distributions	39
4.3 The inverse transmittance problem	41
4.4 Continuously varying refractive index	43
II Implementation of the FDTD numerical method	45
5 Discretizing the Maxwell's equations	47
5.1 Time-domain bulk algorithm	48

5.2	Stability	55
5.3	Frequency domain analysis	57
6	Absorbing boundary conditions	63
6.1	Convolutional-Perfectly Matched Layer	64
6.2	Periodic boundary conditions	81
7	Linearly polarized plane wave excitation	83
7.1	Propagation along a coordinate axis	85
7.2	Propagation along arbitrary directions	86
8	Algorithm modification for dispersive materials	87
8.1	Recursive convolution	87
8.2	Auxiliary differential equation	88
9	FDTD algorithm parallelization	93
9.1	Domain decomposition	94
9.2	Suitable MPI data structures	97
III	Applications of the FDTD method	101
10	Validation of the implemented FDTD method	103
10.1	Field inside and scattered off a sphere	103
10.2	Scalability of the parallelized code	107
11	Plasmonics applications	111
11.1	Scattering from spherical nanoparticles. Analytical case.	112
11.2	Scattering from nonspherical nanoparticles. FDTD case.	118
11.3	Arrayed particles on a silicon substrate.	121
12	Photonics application to an opal crystal	133
12.1	Eigenvalue problem and the band structure	134
12.2	The FCC opal photonic crystal	136
12.3	Multilayer equivalence: homogenization	147
12.4	Dispersion band structure reconstruction	150
A	Some dyadics calculations	155
B	Some Gaussian integrals	157
C	Stability condition for the FDTD method	159

CONTENTS

v

Conclusions

161

Introduction and motivation

The present work analyzes and describes a method for the direct numerical solution of the Maxwell's equations of classical electromagnetism. This is the FDTD (Finite-Difference Time-Domain) method, along with its implementation in an "in-house" computing code for large parallelized simulations. Both are then applied to the modelization of photonic and plasmonic structures interacting with light. These systems are often too complex, either geometrically and materially, in order to be mathematically tractable and an exact analytic solution in closed form, or as a series expansion, cannot be obtained. The only way to gain insight on their physical behavior is thus to try to get a numerical approximated, although convergent, solution.

This is a current trend in modern physics because, apart from perturbative methods and asymptotic analysis, which represent, where applicable, the typical instruments to deal with complex physico-mathematical problems, the only general way to approach such problems is based on the direct approximated numerical solution of the governing equations. Today this last choice is made possible through the enormous and widespread computational capabilities offered by modern computers, in particular High Performance Computing (HPC) done using parallel machines with a large number of CPUs working concurrently. Computer simulations are now a sort of virtual laboratories, which can be rapidly and costless setup to investigate various physical phenomena. Thus computational physics has become a sort of third way between the experimental and theoretical branches.

The plasmonics application of the present work concerns the scattering and absorption analysis from single and arrayed metal nanoparticles, when surface plasmons are excited by an impinging beam of light, to study the radiation distribution inside a silicon substrate behind them. This has potential applications in improving the efficiency of photovoltaic cells.

The photonics application of the present work concerns the analysis of the optical reflectance and transmittance properties of an opal crystal. This is a regular and ordered lattice of macroscopic particles which can stop light propagation in certain wavelength bands, and whose study has potential ap-

plications in the realization of low threshold laser, optical waveguides and sensors. For these latter, in fact, the crystal response is tuned to its structure parameters and symmetry and varies by varying them.

The present work about the FDTD method represents an enhancement of a previous one made for my MSc Degree Thesis in Physics, which has also now geared toward the visible and neighboring parts of the electromagnetic spectrum. It is organized in the following fashion.

Part I provides an exposition of the basic concepts of electromagnetism which constitute the minimum, although partial, theoretical background useful to formulate the physics of the systems here analyzed or to be analyzed in possible further developments of the work. It summarizes Maxwell's equations in matter and the time domain description of temporally dispersive media. It addresses also the plane wave representation of an electromagnetic field distribution, mainly the far field one. The Kirchhoff formula is described and deduced, to calculate the angular radiation distribution around a scatterer. Gaussian beams in the paraxial approximation are also slightly treated, along with their focalization by means of an approximated diffraction formula useful for their numerical FDTD representation. Finally, a thorough description of planarly multilayered media is included, which can play an important ancillary role in the homogenization procedure of a photonic crystal, as described in Part III, but also in other optical analyses.

Part II properly concerns the FDTD numerical method description and implementation. Various aspects of the method are treated which globally contribute to a working and robust overall algorithm. Particular emphasis is given to those arguments representing an enhancement of previous work. These are: the analysis from existing literature of a new class of absorbing boundary conditions, the so called Convolutional-Perfectly Matched Layer, and their implementation; the analysis from existing literature and implementation of the Auxiliary Differential Equation Method for the inclusion of frequency dependent electric permittivity media, according to various and general polarization models; the description and implementation of a "plane wave injector" for representing impinging beam of lights propagating in an arbitrary direction, and which can be used to represent, by superposition, focalized beams; the parallelization of the FDTD numerical method by means of the Message Passing Interface (MPI) which, by using the here proposed, suitable, user defined MPI data structures, results in a robust and scalable code, running on massively parallel High Performance Computing Machines like the IBM/BlueGeneQ with a core number of order 2×10^5 .

Finally, Part III gives the details of the specific plasmonics and photonics applications made with the "in-house" developed FDTD algorithm, to demonstrate its effectiveness. After Chapter 10, devoted to the validation

of the FDTD code implementation against a known solution, Chapter 11 is about plasmonics, with the analytical and numerical study of single and arrayed metal nanoparticles of different shapes and sizes, when surface plasmon are excited on them by a light beam. The presence of a passivating embedding silica layer and a silicon substrate are also included. The next Chapter 12 is about the FDTD modelization of a face-cubic centered (FCC) opal photonic crystal sample, with a comparison between the numerical and experimental transmittance/reflectance behavior. An homogenization procedure is suggested of the lattice discontinuous crystal structure, by means of an averaging procedure and a planarly multilayered media analysis, through which better understand the reflecting characteristic of the crystal sample. Finally, a procedure for the numerical reconstruction of the crystal dispersion banded $\omega - k$ curve inside the first Brillouin zone is proposed. Three Appendices providing details about specific arguments dealt with during the exposition conclude the work.

Part I

Some selected topics in electromagnetism and wave optics

Chapter 1

Electromagnetism in matter

1.1 Maxwell's equations in matter

The electromagnetic field inside matter is described by the Maxwell's equations which, using the SI unit system are:

$$\vec{\nabla} \cdot \vec{D} = \rho \quad (1.1a)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (1.1b)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (1.1c)$$

$$\vec{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{j} \quad (1.1d)$$

where \vec{E} is the *electric* field in Volt/m, \vec{D} is the *electric induction* field in Coulomb/m², \vec{H} is the *magnetic* field in Ampère/m, \vec{B} is the *magnetic induction* in Weber/m². In Eq.n (1.1a) and Eq.n (1.1d), ρ and \vec{j} are the total free electric charge density in Coulomb/m³ and the total free electric current density in Ampère/m², respectively. Here, free means unbounded. They have to satisfy locally the charge conservation law:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad (1.2)$$

i.e., to satisfy a continuity equation. Here the first order vector operator $\vec{\nabla}$:

$$\vec{\nabla} \equiv \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}$$

(\hat{x} , \hat{y} , \hat{z} are the unit vectors of the cartesian orthogonal reference frame) is used to denote, along with the \cdot and \times algebraic formal operations in cartesian coordinates, the divergence ($\vec{\nabla} \cdot$) and curl ($\vec{\nabla} \times$) operators, but obviously Eq.ns (1.1) hold in any coordinate system. To solve the Maxwell's equations, *constitutive relations* have to be specified. The most usual are:

$$\vec{D}(\vec{r}, t) = \epsilon(\vec{r})\vec{E}(\vec{r}, t) \quad (1.3)$$

$$\vec{B}(\vec{r}, t) = \mu(\vec{r})\vec{H}(\vec{r}, t) \quad (1.4)$$

$$\vec{j}_{cond}(\vec{r}, t) = \sigma(\vec{r})\vec{E}(\vec{r}, t) \quad (1.5)$$

where ϵ and μ are the absolute electric permittivity and magnetic permeability of the media, in Farad/m and Henry/m respectively, while σ is the electric conductivity in Siemens/m. Eq.n (1.5) implies that the total free charge and currents are decomposed as:

$$\rho = \rho_{cond} + \rho_{source}$$

$$\vec{j} = \vec{j}_{cond} + \vec{j}_{source}$$

and each contribution verifies independently a continuity equation like (1.2). The source contributions refer to the impressed charges and currents from generators. The remaining free charges and currents constitute the ohmic contribution. Bounded charges and currents are instead taken in account through relations (1.3) and (1.4). A first generalization of the above linear constitutive relations is by means of a tensorial permittivity, or permeability, or conductivity. A further one is by means of non-local temporal and spatial linear relations. In the present work non-magnetic materials are considered throughout, which means (without lack of generality)

$$\mu(\vec{r}) \equiv \mu_o = 4\pi \times 10^{-7} \frac{\text{Henry}}{\text{m}}$$

everywhere, μ_o being the vacuum permeability. Instead, the most general linear relation between \vec{D} and \vec{E} that will be considered in the present work is a temporal, non-local, scalar one:

$$\vec{D}(\vec{r}, t) = \epsilon_o \left[\epsilon_{r,\infty}(\vec{r})\vec{E}(\vec{r}, t) + \int_{-\infty}^{+\infty} G(\vec{r}, \tau)\vec{E}(\vec{r}, t - \tau)d\tau \right] \quad (1.6)$$

where the causality principle imposes the kernel G to be zero for $\tau < 0$, because contributions to \vec{E} in advance with respect to the present time cannot

exist. When translated in the frequency domain Eq.n (1.6) becomes, being the transform of the temporal convolution the product of the trasforms of the convolving functions:

$$\vec{D}(\vec{r}, \omega) = \epsilon_o [\epsilon_{r,\infty}(\vec{r}) + \chi(\vec{r}, \omega)] \vec{E}(\vec{r}, \omega) \quad (1.7)$$

where, with an abuse of notation, the same letters have been used for the time domain electric field and its frequency domain counterpart. The general relation between \vec{D} and \vec{E} is

$$\vec{D} = \epsilon_o \vec{E} + \vec{P} \quad (1.8)$$

where \vec{P} is the matter polarization field in Coulomb/*textm*² (in both the time and frequency domains). By comparing (1.6) and (1.7) with $\epsilon_{r,\infty} = 1$, and (1.8), we have that the electric susceptibility $\chi(\omega)$ — omitting the \vec{r} dependence —, which connects in the linear regime \vec{P} to \vec{E} : $\vec{P} = \epsilon_o \chi \vec{E}$, is given by the transform pair:

$$\chi(\omega) = \int_{-\infty}^{+\infty} G(t) e^{+i\omega t} dt \quad (1.9a)$$

$$G(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \chi(\omega) e^{-i\omega t} d\omega. \quad (1.9b)$$

Note that in the present work the phase factor $e^{-i\omega t}$ has been chosen for the temporal harmonic fields: $E(t) = Re \{E_\omega e^{-i\omega t}\}$. On the other hand, any multiplicative factor in the definition of the transform pair

By comparing (1.3) and (1.6) we see that the former is justified if $G \equiv 0$ and only the instantaneous response is retained, corresponding to the limit $\omega \rightarrow \infty$ (from which the ∞ subscript). Thus relation (1.7) is the more correct one. It is usually rewritten by introducing the absolute complex permittivity ϵ_c (i is the imaginary unit and the plus sign is coherent with the choice of the sign in the time harmonic phase factor)

$$\epsilon_c = \epsilon' + i\epsilon''$$

Sometimes to simplify things the instantaneous conductivity response σ enters ϵ'' by inserting (1.3) and (1.5) in (1.1d) and then Fourier transforming both members, thus getting:

$$\epsilon'' = \frac{\sigma}{\omega}.$$

The determination of the kernel function $G(t)$ in Eq. (1.9) is of relevance, because in a time domain approach such as the one presented in this work, the equation to be discretized is (1.6) and not (1.7).

1.2 Temporally dispersive media.

Various polarization models are used to determine the electric susceptibility function $\chi(\omega)$ and to which, through (1.9b), correspond different kernels $G(t)$. The latter are explicitly calculated by means of a complex contour integration, picking up the residue contributions from the poles of χ in the complex z plane, with $\omega = \text{Re}\{z\}$. In what follows, in the present Section, $\theta(t)$ is the unit step distribution: $\theta(t) = 1$ if $t > 0$, $\theta(t) = 0$ if $t < 0$. It arises from an application of the Jordan lemma of complex analysis to the semicircle path of integration (all the poles lie in the lower complex half plane). It represents the mathematical expression of *causality*: the effect can only follow the cause.

- Drude (or Drude-Sommerfeld) model [18]:

$$\chi_{DS}(\omega) = -\frac{\omega_{\text{pl}}^2}{\omega(\omega + i\gamma)} \quad (1.10)$$

where ω_{pl} is a plasma frequency of the free-electron gas contributing to the conductivity of *metals* and γ is a damping factor (both parameters positive), for which:

$$G_{DS}(t) = \frac{\omega_{\text{pl}}^2}{\gamma} (1 - e^{-\gamma t}) \theta(t).$$

- Lorentz model [18]:

$$\chi_L(\omega) = \frac{(\epsilon_{r,s} - \epsilon_{r,\infty})\omega_o^2}{\omega_o^2 - 2i\omega\gamma - \omega^2} \quad (1.11)$$

where ω_o is a resonance frequency of the *bounded* oscillating electrons and γ their damping factor ($\omega_o > \gamma > 0$). After setting: $\tilde{\omega} = \sqrt{\omega_o^2 - \gamma^2}$, it results that:

$$G_L(t) = \frac{(\epsilon_{r,s} - \epsilon_{r,\infty})\omega_o^2}{\tilde{\omega}} e^{-\gamma t} \sin(\tilde{\omega}t)\theta(t).$$

This model generalizes to a sum of N terms with different resonance frequencies and damping factors like in (8.1) of Section 8.2, Chapter 8. $G_L(t)$ varies accordingly, becoming a sum of N terms.

- Debye model:

1.3. SOLUTION FOR THE SPHERE IN THE FIELD OF A PLANE WAVE 7

$$\chi_D(\omega) = \frac{(\epsilon_{r,s} - \epsilon_{r,\infty})i\gamma}{\omega + i\gamma} \quad (1.12)$$

is like the Drude-Sommerfeld case, but without the simple pole in the origin. It results that:

$$G_D(t) = \gamma e^{-\gamma t} \theta(t).$$

• Critical Points model [48] and References 11–12 therein, is an ad hoc susceptibility of the form:

$$\chi_{CP}(\omega) = A \Omega \left(\frac{e^{i\phi}}{\Omega - \omega - i\Gamma} + \frac{e^{-i\phi}}{\Omega + \omega + i\Gamma} \right) \quad (1.13)$$

where A, Ω, Γ, ϕ are settable real parameters (apart ϕ which can have both signs, all other parameters are positive definite). A *two points* critical model, i.e., two terms like the one above (with parameters indexed with $\ell = 1, 2$ and with suitable values), in conjunction with a Drude model, has been proposed to fit accurately the complex permittivity of noble metals like gold and silver in the wavelength range $200 \div 1000$ nm. In [48] parameter values for aluminum and chromium are also given. It should be observed that the complex permittivity fitting is constrained to expressions obeying appropriate *dispersion relations* [7,8] between their real and imaginary parts. The contribution to $G_{CP}(t)$ from a single term like the one above is:

$$G_{CP}(t) = 2A \Omega e^{-\Gamma t} \sin(\Omega t - \phi) \theta(t).$$

which generalizes to a sum for $\ell = 1, 2$ plus a further $G_{DS}(t)$ term for the full proposed model in [48] and References 11–12 therein.

It can be easily shown, possibly by considering the sin function as a complex exponential and taking the imaginary part at the end of calculations that, for all the kernels $G(t)$ above, the time convolution (1.6), once it has been discretized according to the FDTD method that will be described in the Part II of the present work — see Section 8.1 of Chapter 8 and Section 6.1 of Chapter 6 — can be recursively updated, thus permitting to include, in the ensuing numerical algorithm, temporally dispersive media.

1.3 Solution for the sphere in the field of a plane wave

Analytical closed form solutions of electromagnetic problems are hardly to find, except for some particular configurations. An example of an ana-

lytical solution which is not in closed form but, however, given as a series expansion, is that of a sphere of radius a , made of a non-magnetic material of complex electric permittivity ϵ , in the electromagnetic field of an homogeneous, linearly polarized, monochromatic plane wave of given amplitude, propagating along a given direction, in the vacuum (μ_o, ϵ_o). From a numerical evaluation of the various term of the series, it is possible to sum a finite number of them to get, to a given accuracy — within the accumulation error due to the finite precision arithmetic of a computer —, the distribution of the “exact” solution inside and outside the sphere. This serves for comparison with the solution calculated by the FDTD method (or, eventually, any other method) and for testing the effectiveness and accuracy of the latter.

By introducing a “fixed” cartesian right-handed reference frame with origin at the center of the sphere, its positive y axis in the direction of propagation and its z axis in the direction of the plane wave electric vector yields, after the introduction of spherical coordinates with the colatitude angle θ measured from the positive y axis and the azimuthal angle ϕ measured from the positive z axis:

$$x = r \sin(\theta) \sin(\phi)$$

$$y = r \cos(\theta)$$

$$z = r \sin(\theta) \cos(\phi).$$

The unit vectors of the “moving” frame at (r, θ, ϕ) are $\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi$:

$$\hat{e}_r = \sin(\theta) \sin(\phi) \hat{x} + \cos(\theta) \hat{y} + \sin(\theta) \cos(\phi) \hat{z}$$

$$\hat{e}_\theta = \cos(\theta) \sin(\phi) \hat{x} - \sin(\theta) \hat{y} + \cos(\theta) \cos(\phi) \hat{z}$$

$$\hat{e}_\phi = \cos(\phi) \hat{x} - \sin(\phi) \hat{z}$$

with $\hat{x}, \hat{y}, \hat{z}$ the fixed unit vectors. In time-free form (assuming the usual $e^{-i\omega t}$ factor and, with an abuse of notation, using the same letters for the time-domain and the frequency-domain variable fields vectors) the Maxwell’s curl equations are:

$$\vec{\nabla} \times \vec{E} = i\omega\mu_o\vec{H}$$

$$\vec{\nabla} \times \vec{H} = -i\omega\epsilon_c\vec{E}$$

1.3. SOLUTION FOR THE SPHERE IN THE FIELD OF A PLANE WAVE 9

where:

$$\epsilon_c = \begin{cases} \epsilon_o & \text{outside the sphere} \\ \epsilon + i\frac{\sigma}{\omega} & \text{inside the sphere} \end{cases}$$

(alternatively, inside the sphere ϵ_c could be one of the general expressions given in the preceding Section 1.2 for dispersive media. Note that here is the static conductivity σ to contribute the imaginary part of ϵ_c). Moreover, both \vec{E} and \vec{H} have to be solenoidal:

$$\vec{\nabla} \cdot \vec{E} = \vec{\nabla} \cdot \vec{H} = 0.$$

The above two curl equations can be combined to give a second order vector wave equation for \vec{E} (or \vec{H}):

$$\vec{\nabla} \times \vec{\nabla} \times \vec{E} = \omega^2 \mu_o \epsilon_c \vec{E}.$$

Representing the vector fields by means of the $\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi$ base does not allow to write three simple scalar wave equations (or Helmholtz equations) for each one of the components of \vec{E} , because those versors are not spatially constant. In [5] it is shown that if ψ is a scalar solution of the Helmholtz equation:

$$(\vec{\nabla}^2 + k^2)\psi = 0$$

where $k = k_c = \frac{\omega}{c_o} \sqrt{\epsilon_c/\epsilon_o}$ (a complex quantity) inside the sphere, or $k = k_o$ (a real quantity) outside, with c_o the vacuum light speed. Then:

$$\vec{L} = \vec{\nabla} \psi$$

$$\vec{M} = \vec{\nabla} \times (\hat{e}_r r \psi) = \vec{L} \times \vec{r}$$

$$\vec{N} = \frac{1}{k} \vec{\nabla} \times \vec{M}$$

($\vec{r} = r\hat{e}_r$) are three vectorially independent solutions of the vector wave equation. In particular \vec{M} and \vec{N} are solenoidal, thus they are the solutions of interest in the present Section. They can be constructed starting from the solutions ψ of the Helmholtz equation, which is separable in the spherical coordinates r, θ, ϕ . A complete set of solutions is [5]:

$$\psi_{m,n}^{e,o} = z_m(kr) P_m^n(\cos\theta) f^{e,o}(n\phi)$$

where:

- $m = 0, 1, 2, \dots$;

- $n = 0, 1, \dots, m$;
- e, o superscripts stand for *even* or *odd* respectively with:

$$f^e(n\phi) = \cos(n\phi)$$

$$f^o(n\phi) = \sin(n\phi);$$

- $P_m^n(x)$ ($0 \leq x \leq 1$) are associated Legendre functions of the first kind;
- $z_m(\zeta)$, $\zeta = kr$, are spherical Bessel functions:

$$z_m(\zeta) = \frac{Z_m(\zeta)}{\zeta^{\frac{1}{2}}}$$

where $Z_m(\zeta)$ are half-integer order Bessel functions. The spherical Bessel functions are of three kinds, depending on their asymptotic behavior for $\zeta \rightarrow 0$ and $\zeta \rightarrow \infty$ (the point at infinity because, in general, ζ is a complex quantity).

- First kind ($z_m = j_m(\zeta)$):

$$j_m(\zeta) \sim \begin{cases} \zeta^m & \zeta \rightarrow 0 \\ \frac{\sin(\zeta - \frac{m\pi}{2})}{\zeta} & |\zeta| \gg 1, m \end{cases}$$

- Second kind or Neumann ($z_m = n_m(\zeta)$):

$$n_m(\zeta) \sim \begin{cases} \frac{1}{\zeta^{m+1}} & \zeta \rightarrow 0 \\ -\frac{\cos(\zeta - \frac{m\pi}{2})}{\zeta} & |\zeta| \gg 1, m \end{cases}$$

- Third kind or Hankel ($z_m = h_m^{(\pm)}(\zeta) = j_m(\zeta) \pm in_m(\zeta)$):

$$h_m^{(\pm)}(\zeta) \sim (\mp i)^{m+1} \frac{e^{\pm i\zeta}}{\zeta} \quad |\zeta| \gg 1, m.$$

$h_m^{(\pm)}(\zeta)$ represent outgoing/ingoing traveling waves (with respect to the origin of the coordinates).

In $\psi_{m,n}^{e,o}$ the angular dependence is kept separated between even and odd functions, where usually in the spherical harmonic functions they are merged in a single imaginary exponential with $m = -n, \dots, n$. By inserting in the above expressions for \vec{M} and \vec{N} yields the *vector spherical functions*, which form a basis for expanding the solenoidal solutions of the vector wave equation. Explicitly:

1.3. SOLUTION FOR THE SPHERE IN THE FIELD OF A PLANE WAVE11

$$\begin{aligned}\vec{M}_{m,n}^{e,o} = & \frac{1}{\sin(\theta)} z_m(kr) P_m^n(\cos\theta) \frac{df^{e,o}(n\phi)}{d\phi} \hat{e}_\theta + \\ & - z_m(kr) \frac{dP_m^n(\cos\theta)}{d\theta} f^{e,o}(n\phi) \hat{e}_\phi\end{aligned}$$

$$\begin{aligned}\vec{N}_{m,n}^{e,o} = & \frac{m(m+1)}{kr} z_m(kr) P_m^n(\cos\theta) f^{e,o}(n\phi) \hat{e}_r + \\ & + \frac{1}{kr} \frac{d[rz_m(kr)]}{dr} \frac{dP_m^n(\cos\theta)}{d\theta} f^{e,o}(n\phi) \hat{e}_\theta + \\ & + \frac{1}{kr \sin(\theta)} \frac{d[rz_m(kr)]}{dr} \frac{dP_m^n(\cos\theta)}{d\theta} \frac{f^{e,o}(n\phi)}{d\phi} \hat{e}_\phi.\end{aligned}$$

for $m = 0, 1, 2, \dots$ and $n = 0, 1, \dots, m$. Now, inside the sphere the electric field is \vec{E}_t , a “transmitted” one. Outside, it is the vectorial sum of the incident and the “reflected” ones: $\vec{E}_i + \vec{E}_r$. Similarly for \vec{H} . The boundary condition at the spherical interface $r = a$ is the continuity of the tangential field components, of both \vec{E} and \vec{H} :

$$\begin{aligned}\left[\hat{e}_r \times \left(\vec{E}_i + \vec{E}_r \right) \right]_{r=a} &= \left[\hat{e}_r \times \vec{E}_t \right]_{r=a} \\ \left[\hat{e}_r \times \left(\vec{H}_i + \vec{H}_r \right) \right]_{r=a} &= \left[\hat{e}_r \times \vec{H}_t \right]_{r=a}.\end{aligned}$$

The incident, monochromatic, linearly polarized plane wave with a normalized amplitude is:

$$\vec{E}_i = \hat{z} e^{ik_o y} = \hat{z} e^{ik_o r \cos(\theta)} = \sum_{m=1}^{\infty} \left(a_m \vec{M}_{m,1}^{o,I} - i b_m \vec{N}_{m,1}^{e,I} \right)$$

(see [5]) where the choice between even and odd in the expansion is suggested by comparison of the above expressions for \vec{M}_m and \vec{N}_m with the dependence of \hat{z} on the ϕ angle when it is represented in the $\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi$ basis. For the same reason there is no sum on the n index, which is fixed at 1. The superscript I indicates that as spherical Bessel functions are chosen those of the first kind, which are regular at the origin. Also note that $k = k_o$ (also in \vec{M}_m and \vec{N}_m), i.e., that of the vacuum. The expansion coefficients a_m and b_m are found using the orthogonality of the base functions and are [5]:

$$a_m = b_m = \frac{2m+1}{m(m+1)} i^m.$$

The incident magnetic field, being directed along \hat{x} is represented as:

$$\vec{H}_i = \hat{x} \frac{k_o}{\omega\mu_o} e^{ik_o y} = -\frac{k_o}{\omega\mu_o} \sum_{m=1}^{\infty} \left(b_m \vec{M}_{m,1}^{e,I} + ia_m \vec{N}_{m,1}^{o,I} \right)$$

this expression being again dictated by the dependence of \hat{x} on the ϕ angle when it is represented in the $\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi$ basis. The “reflected” and “transmitted” fields can be expanded by similarity with the above expressions in the following way:

$$\vec{E}_r = \sum_{m=1}^{\infty} i^m \frac{2m+1}{m(m+1)} \left(a_m^r \vec{M}_{m,1}^{o,III} - ib_m^r \vec{N}_{m,1}^{e,III} \right) \quad (1.14a)$$

$$\vec{H}_r = -\frac{k_o}{\omega\mu_o} \sum_{m=1}^{\infty} i^m \frac{2m+1}{m(m+1)} \left(b_m^r \vec{M}_{m,1}^{e,III} + ia_m^r \vec{N}_{m,1}^{o,III} \right) \quad (1.14b)$$

where the spherical Bessel functions of the third kind are used for the “reflected” field, because they have the correct behavior at a great distance from the sphere and satisfy the Sommerfeld radiation condition. And for the field inside the sphere:

$$\vec{E}_t = \sum_{m=1}^{\infty} i^m \frac{2m+1}{m(m+1)} \left(a_m^t \vec{M}_{m,1}^{o,I} - ib_m^t \vec{N}_{m,1}^{e,I} \right)$$

$$\vec{H}_t = -\frac{k_c}{\omega\mu_o} \sum_{m=1}^{\infty} i^m \frac{2m+1}{m(m+1)} \left(b_m^t \vec{M}_{m,1}^{e,I} + ia_m^t \vec{N}_{m,1}^{o,I} \right).$$

Here the complex k_c and the spherical Bessel functions of the first kind have to be used in \vec{M}_m and \vec{N}_m because they are regular at the origin.

The four unknown coefficients, for each $m = 1, 2, \dots$, in the above expansions: $a_m^r, b_m^r, a_m^t, b_m^t$, are found by imposing the boundary conditions at $r = a$ for $E_\theta, E_\phi, H_\theta, H_\phi$ and equating term by term the expansions. This gives:

$$a_m^r = -\frac{j_m(k_c a)[k_o a j_m(k_o a)]' - j_m(k_o a)[k_c a j_m(k_c a)]'}{j_m(k_c a)[k_o a h_m(k_o a)]' - h_m(k_o a)[k_c a j_m(k_c a)]'} \quad (1.15a)$$

$$b_m^r = -\frac{j_m(k_o a)[k_c a j_m(k_c a)]' - \beta^2 j_m(k_c a)[k_o a j_m(k_o a)]'}{h_m(k_o a)[k_c a j_m(k_c a)]' - \beta^2 j_m(k_c a)[k_o a h_m(k_o a)]'} \quad (1.15b)$$

$$a_m^t = -\frac{j_m(k_o a)[k_o a h_m(k_o a)]' - h_m(k_o a)[k_o a j_m(k_o a)]'}{h_m(k_o a)[k_c a j_m(k_c a)]' - j_m(k_c a)[k_o a h_m(k_o a)]'} \quad (1.15c)$$

$$b_m^t = \beta \frac{h_m(k_o a)[k_o a j_m(k_o a)]' - j_m(k_o a)[k_o a h_m(k_o a)]'}{h_m(k_o a)[k_c a j_m(k_c a)]' - \beta^2 j_m(k_c a)[k_o a h_m(k_o a)]'} \quad (1.15d)$$

where:

$$\beta = \frac{k_c}{k_o}$$

and

$$[k a z_m(k a)]' \equiv \left. \frac{d[\zeta z_m(\zeta)]}{d\zeta} \right|_{\zeta=k a}.$$

Moreover, the Hankel functions used are effectively $h_m^{(+)}$, i.e., those corresponding to outgoing waves (the superscript has been omitted in the expressions above for the coefficients for better readability).

A computer code to calculate numerically the fields $\vec{E}_r, \vec{E}_t, \vec{H}_r, \vec{H}_t$ (\vec{E}_i and \vec{H}_i , which are to be added outside the sphere to \vec{E}_r and \vec{H}_r , do not need to be calculated through their expansions, because are immediately known from their imaginary exponential form $\propto e^{i k_o \cos(\theta)}$) has to be able to evaluate the required spherical Bessel functions of complex argument and the associated Legendre functions, as well as their first derivatives. Once the eabove expressions for the four kinds of expansion coefficients have been evaluated at a given (angular) frequency ω , for the given sphere of radius a , with the given ϵ and σ parameters and for a whole range of $m = 1, \dots, m_{max}$ indices, they can be inserted in the above expressions for the field expansions and by summing up the contributions from the m_{max} terms yields the field components (which can eventually be transformed in the cartesian ones) values at any chosen (r, θ, ϕ) point. It is expected that the more terms in the expansions (increasing m_{max}), the more accurate their numerically calculated values, even if the more longer the calculations. As with other generalized Fourier expansions, if the represented fields are sufficiently smooth, besides the mean-square convergence, the expansions converge also in the ordinary sense. The evaluation of special functions and their derivatives requires dedicated mathematical numerical libraries to be linked to the computer code (or the use of interpreted languages like Python, which has a dedicated module implementing such special function but which is unavoidably much slower than a compiled code). Alternatively, they can be calculated recursively, exploiting the recurrence relations they satisfy, or by means of their series expansions, as explained in [40].

The above described analytical procedure (and the related numerical implementation) could be generalized further, within a spherically symmetric context, to treat a layered sphere made of a finite number of concentric spherical shells, each one made of a different material of given complex permittivity. Boundary conditions would have to be imposed at each spherical interface corresponding to the transition between different media.

Chapter 2

Spatial Fourier analysis and far-fields

In the present Chapter the electromagnetic field vectors \vec{E} and \vec{H} depend on the position but also on the (angular) frequency ω , which is not explicitly indicated. With an abuse of notation the same symbols are used here for frequency domain fields as were used in Chapter 1 for the time domain ones.

2.1 Plane wave spectrum representation

Once the spatial distribution \vec{E} of the electric field has been determined at a given angular frequency ω in a given region of a homogeneous medium free of sources (\vec{E} will be in general a complex vector), one can consider a given $y = \text{const.}$ plane and make a two-dimensional spatial direct/inverse Fourier transform:

$$\vec{\mathcal{E}}(k_x, k_z; y) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dz \vec{E}(x, y, z) e^{-i(k_x x + k_z z)} \quad (2.1a)$$

$$\vec{E}(x, y, z) = \int_{-\infty}^{+\infty} dk_x \int_{-\infty}^{+\infty} dk_z \vec{\mathcal{E}}(k_x, k_z; y) e^{+i(k_x x + k_z z)}. \quad (2.1b)$$

Both the above Fourier integrals hold separately for each vector component of $\vec{\mathcal{E}}$ and \vec{E} . Because the left hand member of (2.1b) has to be the solution of the scalar Helmholtz equation:

$$\left(\vec{\nabla}^2 + k^2\right) \vec{E}(\vec{r}) = \vec{0}$$

it is easily seen that as a consequence $\vec{\mathcal{E}}$ satisfies:

$$\left[\frac{\partial^2}{\partial y^2} + (k^2 - k_x^2 - k_z^2) \right] \vec{\mathcal{E}}(k_x, k_z; y) = \vec{0}.$$

Taking the initial condition at $y = 0$, the *object plane*, one has:

$$\vec{\mathcal{E}}(k_x, k_z; y) = \vec{\mathcal{E}}(k_x, k_z; 0) e^{\pm i k_y y}$$

with

$$k_y = \sqrt{k^2 - k_x^2 - k_z^2},$$

being $k = \omega/c$ with c the light speed in the given homogeneous medium. Choosing for definitness the forward y direction (i.e., the plus sign in the exponential) and $Im\{k_y\} > 0$ (i.e., the positive branch of the square root) in such a way that, for $y \rightarrow +\infty$, the field remains finite, (2.1b) can be rewritten as:

$$\vec{E}(x, y, z) = \int_{-\infty}^{+\infty} dk_x \int_{-\infty}^{+\infty} dk_z \vec{\mathcal{E}}(k_x, k_z; 0) e^{+i(k_x x + k_y y + k_z z)}. \quad (2.2)$$

Thus at any given *image plane* $y = \text{const.}$ the field can be reconstructed if its spectrum is known in the object plane. (2.2) is known as the *angular spectrum representation* of \vec{E} because it is an integral sum over a set of plane waves propagating in various angularly distributed directions, weighted by $\vec{\mathcal{E}}$. However, when the independent variables are such that:

$$k_x^2 + k_z^2 > k^2$$

there are evanescent non-homogeneous plane waves, decaying exponentially with increasing y and oscillating sinusoidally in transverse directions. The solenoidality condition must also be imposed on the spectrum $\vec{\mathcal{E}}$:

$$\vec{k} \cdot \vec{\mathcal{E}} = 0$$

for it can correctly describe a solution of Maxwell's equations. Similar conclusions can be applied to the magnetic spectrum $\vec{\mathcal{H}}$, which can be calculated from $\vec{\mathcal{E}}$ using the $\vec{\nabla} \times$ operator.

2.2 Far-field in the angular spectrum representation

An important asymptotic analysis can be performed on (2.2) by means of the *stationary phase* method [4, 18], in evaluating the far-zone approximation for $\vec{E}(\vec{r})$ at an infinite distance from the $y = 0$ object plane, in the $y \geq 0$ half-space. Thus, by starting from (2.2) and introducing the unit vector \hat{s} to specify directions:

$$\hat{s} = (s_x, s_y, s_z) = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r} \right)$$

($r = \|\vec{r}\| = \sqrt{x^2 + y^2 + z^2}$), by taking the limit $r \rightarrow \infty$ one can write:

$$\vec{E}_\infty(x, y, z) = \lim_{kr \rightarrow \infty} \iint_{k_x^2 + k_z^2 \leq k^2} \vec{\mathcal{E}}(k_x, k_z; 0) e^{+ikr(\frac{k_x}{k}s_x + \frac{k_y}{k}s_y + \frac{k_z}{k}s_z)} dk_x dk_z \quad (2.3)$$

in which the contribution of the evanescent waves is a priori neglected because they have an exponential vanishing decay at infinity. To calculate the limiting behavior for $kr \rightarrow \infty$ of (2.3) — with s_x , s_z and $s_y = \sqrt{1 - s_x^2 - s_z^2}$ kept fixed — it is rewritten as:

$$\vec{E}_\infty(x, y, z) = \lim_{\kappa \rightarrow \infty} \iint_{p^2 + q^2 \leq 1} \vec{e}(p, q) e^{+i\kappa(ps_x + ms_y + qs_z)} dp dq, \quad (2.4)$$

where $\kappa = kr$ and $m = \sqrt{1 - p^2 - q^2}$ ($p = \frac{k_x}{k}$, $q = \frac{k_z}{k}$, $m = \frac{k_y}{k}$). One then has to consider the stationary points of the phase function $g(p, q) = ps_x + ms_y + qs_z$ inside the integration domain, which is the unit circle in the pq -plane. If (p', q') is such a stationary point, the final asymptotic result, as given in Subsection 3.3.4 of [4], is:

$$\vec{E}_\infty(s_x r, s_y r, s_z r) \sim \frac{2\pi i \sigma}{kr \sqrt{|\Delta|}} \vec{e}(p', q') e^{ikr(p's_x + m's_y + q's_z)}, \quad (2.5)$$

where $m' = \sqrt{1 - p'^2 - q'^2}$, Δ is the Hessian determinant of $g(p, q)$ at (p', q') and σ depends on the trace Σ of the Hessian matrix evaluated at (p', q') :

$$\sigma = \begin{cases} +1 & \text{if } \Delta > 0, \Sigma > 0 \\ -1 & \text{if } \Delta > 0, \Sigma < 0 \\ -i & \text{if } \Delta < 0 \end{cases} .$$

In the present case for the points where $g(p, q)$ is stationary, one finds that:

$$\frac{p'}{m'} = \frac{s_x}{s_y}, \quad \frac{q'}{m'} = \frac{s_z}{s_y}$$

(s_x , s_y and s_z constant) from which it follows that the solutions for the stationary point are:

$$\boxed{p' = s_x, \quad q' = s_z, \quad m' = s_y} \quad (2.6)$$

which also imply that $g(p', q') = 1$. The physical significance of (2.6) is that *one and only one plane wave of the entire angular spectrum contributes to the far field at a point located in a given direction: namely the wave that propagates in that particular direction, the effects of the other waves canceling each other by destructive interference*. Moreover, resulting $\Sigma < 0$ and

$$\Delta = \frac{1}{s_y^2},$$

one has that $\sigma = -1$ and finally, from (2.5) and (2.6):

$$\vec{E}_\infty(s_x r, s_y r, s_z r) = -2\pi i k_y \vec{\mathcal{E}}(k s_x, k s_z; 0) \frac{e^{i k r}}{r}. \quad (2.7)$$

In getting (2.7) one has to consider that, besides being $k_x = k s_x$, $k_y = k s_y$ and $k_z = k s_z$, $\vec{\mathcal{E}}$ incorporates a factor of k^2 which is missing in \vec{e} , due to the different integration variables in (2.3) and (2.4). Inverting (2.7) emphasizing the k_x , k_z and k_y dependence of $\vec{E}_\infty(s_x r, s_y r, s_z r)$ through s_x , s_z and s_y and using it in (2.2) assuming that only non-evanescent (homogeneous) waves are present, it is possible to express a given field in terms of its far-field plane waves [18]:

$$\boxed{\vec{E}(x, y, z) = \frac{i r e^{-i k r}}{2\pi} \iint_{k_x^2 + k_z^2 \leq k^2} \vec{E}_\infty(k_x, k_z) e^{+i(k_x x + k_y y + k_z z)} \frac{1}{k_y} dk_x dk_z.} \quad (2.8)$$

The approximation $k_y \approx k$ would make (2.8) an exact Fourier transform. It is the Fourier Optics limit.

2.3 Kirchhoff formula and the near to far field transform

The Kirchhoff integral formula is the mathematical expression of Huygens' principle, showing how a wave field on a surface Σ determines the wave

2.3. KIRCHHOFF FORMULA AND THE NEAR TO FAR FIELD TRANSFORM 19

field off the surface Σ . The sketch in Fig. 2.1 below illustrates the situation.

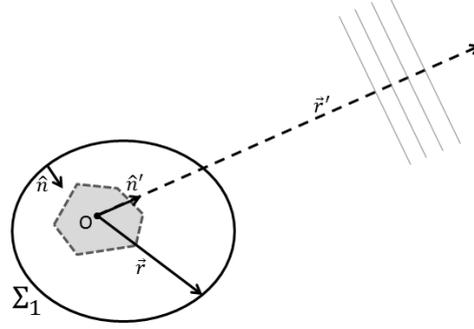


Figure 2.1: Kirchhoff integral formula allows to extrapolate the plane wave far field ($\vec{r}' \rightarrow \infty$) from the knowledge of the field near primary or secondary sources (grey area enclosed inside the integration surface Σ_1).

A surface $\Sigma = \Sigma_1 \cup \Sigma_2$ encloses a region of space V free of sources, which are all contained in the volume inside Σ_1 (the grey area bounded by the shaded line). The final calculations are extrapolated in the limit $\Sigma_2 \rightarrow \Sigma_\infty$, a spherical surface going at infinity and centered around an arbitrary fixed origin O inside Σ_1 . The asymptotic behavior of the wave field, which becomes that of a plane wave, is analyzed along a direction exiting from O and defined by the versor \hat{n}' . The versor \hat{n} instead, indicates the outward normal to Σ . To derive the Kirchhoff formula one can start from the Green's function of the *scalar* Helmholtz equation in an homogeneous medium:

$$\left(\vec{\nabla}^2 + k^2\right) g(\vec{r} - \vec{r}') = -\delta(\vec{r} - \vec{r}')$$

where $k = \omega/c$ (c is the light propagation velocity). The right hand side member represent a point source at \vec{r}' . It is well known [2, 5, 6] that the *fundamental solution* with spherical symmetry around \vec{r}' for a homogeneous unbounded medium is:

$$g(\vec{r} - \vec{r}') = \frac{e^{ik\|\vec{r}-\vec{r}'\|}}{4\pi \|\vec{r} - \vec{r}'\|}$$

where the sign in the exponential is coherent with the choice for the time factor $e^{-i\omega t}$ and represents outgoing waves. By considering now the *vector* Helmholtz equation in a source-free region, obtained from the Maxwell's equations (1.1) expressed in the frequency domain:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{E}(\vec{r}) - k^2 \vec{E}(\vec{r}) = \vec{0}, \quad (2.9)$$

its fundamental solution (see below) in an unbounded homogeneous medium is the *dyadic* Green's function:

$$\bar{\bar{G}}(\vec{r} - \vec{r}') = \left[\bar{\bar{I}} + \frac{\vec{\nabla}' \vec{\nabla}'}{k^2} \right] g(\vec{r} - \vec{r}'), \quad (2.10)$$

a 3×3 matrix generalization of a vector whose components transform like an \mathbb{R}^3 second rank tensor and $\vec{\nabla}' \vec{\nabla}'$ is the dyadic second order differential operator:

$$\vec{\nabla}' \vec{\nabla}' = \sum_{\alpha} \sum_{\beta} \hat{\alpha} \hat{\beta} \frac{\partial^2}{\partial x'_{\alpha} \partial x'_{\beta}}$$

with $\alpha, \beta = x, y, z$ while $\bar{\bar{I}}$ is the unit dyadic:

$$\bar{\bar{I}} = \hat{x} \hat{x} + \hat{y} \hat{y} + \hat{z} \hat{z} = \sum_{\alpha} \sum_{\beta} \delta_{\alpha\beta} \hat{\alpha} \hat{\beta},$$

\hat{x} , \hat{y} and \hat{z} being the cartesian orthonormalized basis. Explicitly:

$$\bar{\bar{G}} = \sum_{\alpha} \sum_{\beta} G_{\alpha\beta} \hat{\alpha} \hat{\beta}$$

where:

$$G_{\alpha,\beta} = g \delta_{\alpha\beta} + \frac{1}{k^2} \frac{\partial^2 g}{\partial x'_{\alpha} \partial x'_{\beta}}. \quad (2.11)$$

Due to the $\vec{r} - \vec{r}'$ dependence of g one has that:

$$\frac{\partial^2 g}{\partial x'_{\alpha} \partial x'_{\beta}} = \frac{\partial^2 g}{\partial x_{\alpha} \partial x_{\beta}}.$$

Both g and $\bar{\bar{G}}$ for homogeneous unbounded media are also symmetric in the exchange of arguments \vec{r} , \vec{r}' . It can then be seen that:

$$\vec{\nabla} \times \vec{\nabla} \times \bar{\bar{G}}(\vec{r} - \vec{r}') - k^2 \bar{\bar{G}}(\vec{r} - \vec{r}') = \bar{\bar{I}} \delta(\vec{r} - \vec{r}'). \quad (2.12)$$

In fact:

$$\vec{\nabla} \times \bar{\bar{G}} \stackrel{\text{def}}{=} \sum_{\alpha} \sum_{\beta} \sum_{\gamma} \partial_{\gamma} G_{\alpha\beta} (\hat{\gamma} \times \hat{\alpha}) \hat{\beta}$$

where, as usual, $\alpha, \beta, \gamma = x, y, z$. Applying again the $\vec{\nabla} \times$ operator to the left of both members of the previous expression and using (2.11), one gets (the details are found in Appendix A):

$$\vec{\nabla} \times \vec{\nabla} \times \bar{G} = \sum_{\alpha} \sum_{\beta} (\partial_{\alpha} \partial_{\beta} g) \hat{\alpha} \hat{\beta} - \sum_{\alpha} (\vec{\nabla}^2 g) \hat{\alpha} \hat{\alpha}$$

and the result follows easily when subtracting the $k^2 \bar{G}$ term and remembering that g is the fundamental solution of the scalar Helmholtz equation. By calculations similar to those in Appendix A, one can also see that an alternative form of (2.10) is:

$$\bar{G}(\vec{r} - \vec{r}') = \frac{1}{k^2} \left[\vec{\nabla} \times \vec{\nabla} \times \bar{I}g(\vec{r} - \vec{r}') - \bar{I}\delta(\vec{r} - \vec{r}') \right], \quad (2.13)$$

from which it also follows that:

$$\vec{\nabla} \times \bar{G}(\vec{r} - \vec{r}') = \vec{\nabla} \times \bar{I}g(\vec{r} - \vec{r}'). \quad (2.14)$$

Now, if \vec{F} and \bar{D} are a vector and a dyadic field respectively, it can be shown (see Appendix A) that:

$$\vec{\nabla} \cdot (\vec{F} \times \bar{D}) = (\vec{\nabla} \times \vec{F}) \cdot \bar{D} - \vec{F} \cdot (\vec{\nabla} \times \bar{D}). \quad (2.15)$$

Moreover, after pre-multiplying (2.12) by $\vec{E}(\vec{r})$, post-multiplying (2.9) by \bar{G} , subtracting the resultant equation from the first one and integrating with respect to \vec{r} over the volume V enclosed by Σ , one has:

$$\vec{E}(\vec{r}') = \int_V dV \left[\vec{E}(\vec{r}) \cdot \vec{\nabla} \times \vec{\nabla} \times \bar{G}(\vec{r} - \vec{r}') - \vec{\nabla} \times \vec{\nabla} \times \vec{E} \cdot \bar{G}(\vec{r} - \vec{r}') \right]$$

By means of (2.15) it can be shown that the integrand in the above equation equals:

$$-\vec{\nabla} \cdot \left[\vec{E}(\vec{r}) \times \vec{\nabla} \times \bar{G}(\vec{r} - \vec{r}') + \vec{\nabla} \times \vec{E} \times \bar{G}(\vec{r} - \vec{r}') \right]$$

which after consideration of the (frequency domain) Maxwell's equation $\vec{H} = (\vec{\nabla} \times \vec{E})/i\omega\mu$ and the dyadic identity: $\vec{A} \cdot (\vec{B} \times \bar{D}) = (\vec{A} \times \vec{B}) \cdot \bar{D}$, allows to write the volume integral as the surface integral:

$$\vec{E}(\vec{r}') = - \oint_{\Sigma} dS \left[\hat{n} \times \vec{E}(\vec{r}) \cdot \vec{\nabla} \times \bar{G}(\vec{r} - \vec{r}') + i\omega\mu \hat{n} \times \vec{H}(\vec{r}) \cdot \bar{G}(\vec{r} - \vec{r}') \right],$$

where $\Sigma = \Sigma_1 \cup \Sigma_2$ is the surface bounding the volume V and \hat{n} is its outward normal versor. In the above expression, the representation (2.13) and the relation (2.14) are now used. Being \vec{r} on S always different from \vec{r}' ,

the Dirac delta does not contribute at all. The last steps comprise the limit $\Sigma_2 \rightarrow \infty$ with \vec{r}' fixed, then the limit $r' \gg r$. Because all field components ψ will end up obeying the Sommerfeld radiation condition [5]:

$$r \left(\frac{\partial \psi}{\partial r} - ik\psi \right) \rightarrow 0$$

for $r \rightarrow \infty$, they will look like an outgoing plane wave on Σ_∞ (for g this is directly seen from its explicit expression above), thus the integrands are $o(1/r^2)$ in this limit and the integral on Σ_∞ , assumed as a spherical surface of radius $r = \infty$ centered at O , vanishes. On the other hand, an asymptotic expression for g when $r' \gg r$ is:

$$g \approx \frac{e^{ikr'}}{4\pi r'} e^{-i\vec{k}' \cdot \vec{r}}$$

where $\vec{k}' = k\hat{n}'$. From this it follows that:

$$\vec{\nabla} g = -i\vec{k}' g.$$

Thus, if \vec{F} is a vector field, one has:

$$\vec{F} \cdot \vec{\nabla} \times \vec{I} g = ig\vec{k}' \times \vec{F}$$

and

$$\vec{F} \cdot \vec{\nabla} \times \vec{\nabla} \times \vec{I} g = ig\vec{k}' \times (\vec{k}' \times \vec{F}).$$

By using $\vec{F} = \hat{n} \times \vec{E}$ or $\vec{F} = \hat{n} \times \vec{H}$, one can put the surface integral in the final form:

$$\vec{E}(\vec{r}') = \frac{e^{ikr'}}{4\pi i\vec{r}'} \vec{k}' \times \oint_{\Sigma_1} dS e^{-ik \cdot \vec{r}} \left\{ \sqrt{\frac{\mu}{\epsilon}} \left[\hat{n}' \times (\hat{n} \times \vec{H}(\vec{r})) \right] - \hat{n} \times \vec{E}(\vec{r}) \right\}. \quad (2.16)$$

This is the Kirchhoff integral formula [6], expressing the transverse radiation field far from primary or secondary sources as a function of the angles θ , ϕ from which the \hat{n}' versor (defining the direction of observation) depends. This radiation amplitude formula allows to calculate the radiation pattern of given sources after the near \vec{E} and \vec{H} fields have been calculated. Formula (2.16) can be implemented in a computer code by discretizing and numerically calculating the surface integral and by sampling the 4π solid angle around the source with finite increments for the $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$ angles.

Chapter 3

Gaussian beams

In the present Chapter \vec{E} and \vec{H} denote *frequency domain* vector spatial distributions in which the dependence from the (angular) frequency ω is not explicitly indicated. With $\vec{\mathcal{E}}(k_x, k_z; y)$, $\vec{\mathcal{H}}(k_x, k_z; y)$ are instead denoted the *spatial Fourier transforms* as introduced and described in the previous Chapter 2. The “main” propagation direction is assumed to be the y -axis.

3.1 Paraxial approximation

Often the light wavefield in an optical system propagates along a certain direction y while spreading only slowly in the transverse direction contained in the xz plane. From a quantitative viewpoint this means that, if $\vec{k} = (k_x, k_y, k_z)$ is the wavevector in the field angular spectrum representation, there is a dominance of k_y over k_x and k_z [18]:

$$k_y = \sqrt{k^2 - k_x^2 - k_z^2} = k\sqrt{1 - \frac{(k_x^2 + k_z^2)}{k^2}} \approx k - \frac{(k_x^2 + k_z^2)}{2k} \quad (3.1)$$

with $k = \|\vec{k}\|$. All this means also that in the scalar Helmholtz equation in free space, the solutions are assumed as: $\psi(x, y, z) = u(x, y, z)e^{ik_y y}$, and the second derivative of $u(x, y, z)$ along the y -axis is ignored compared to the other second derivatives, in such a way that the paraxial Helmholtz equation becomes [17]:

$$\vec{\nabla}_t^2 u + 2ik \frac{\partial u}{\partial y} = 0$$

where $\vec{\nabla}_t^2$ is the *transverse Laplacian* ($\partial^2/\partial x^2 + \partial^2/\partial z^2$) and $k = \omega/c$ with c the light speed in the homogeneous source free medium considered.

3.2 Gaussian laser beams

To describe a laser beam in its fundamental mode, instead of directly trying a solution of the paraxial Helmholtz equation as in [17], it will be considered an electric field spatial gaussian distribution [18] in the $y = 0$ plane, the so called focal plane, given by:

$$\vec{E}(x', 0, z') = \vec{E}_o e^{-\frac{x'^2+z'^2}{w_o^2}},$$

where \vec{E}_o is a constant vector in the transverse xz plane, that will be spatially transformed at $y = 0$:

$$\begin{aligned} \vec{\mathcal{E}}(k_x, k_z; 0) &= \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} dx' \int_{-\infty}^{+\infty} dz' \vec{E}_o e^{-\frac{x'^2+z'^2}{w_o^2}} e^{-i(k_x x' + k_z z')} \\ &= \vec{E}_o \frac{w_o^2}{4\pi} e^{-(k_x^2 + k_z^2) \frac{w_o^2}{4}} \end{aligned}$$

(for the evaluation of the definite integrals see Appendix B). Now, the above expression is used as the angular spectrum in the representation (2.2) of the frequency domain field with (3.1) as k_y :

$$\vec{E}(x, y, z) = \vec{E}_o \frac{w_o^2}{4\pi} e^{iky} \int_{-\infty}^{+\infty} dk_x \int_{-\infty}^{+\infty} dk_z e^{-(k_x^2 + k_z^2) \left(\frac{w_o^2}{4} + i \frac{y}{2k} \right)} e^{i(k_x x + k_z z)},$$

which can be integrated in nearly the same manner as above (see Appendix B) to give the final expression for the paraxial approximation of the vector field in the frequency domain:

$$\vec{E}(x, y, z) = \vec{E}_o \frac{e^{iky}}{\left(1 + i \frac{2y}{kw_o^2}\right)} e^{-\frac{(x^2+z^2)}{w_o^2} \frac{1}{1 + i \frac{2y}{kw_o^2}}}. \quad (3.2)$$

It should be kept in mind that (3.2) is an approximate solution which does not obey Maxwell's equations, even if it represents in the $y = 0$ plane the most realistic achievement of a linearly polarized electromagnetic plane wave. There are also higher order laser modes characterized by different patterns in the focal plane \vec{E} field distribution [18]. The approximation error inherent in (3.1), (3.2) becomes larger the smaller the "beam waist" radius w_o is, when compared to the wavelength in the given medium.

3.3 Focused beams

This section concerns the mathematical description of the electromagnetic field of an optically focalized light beam propagating along a given aa' axis. The focalization is made by means of an aplanatic convergent lens having aa' as its optical axis and focal length f . The description given here is an adapted and partially modified version from [18,41,42]. Aplanatic (= free of spherical aberration) means that all the in-axis coming rays (from the left, say) converge to the focus F being bended (refracted) in correspondence of a sphere of radius f centered at F behind (to the right) the lens (Gaussian reference sphere). See Fig. 3.1.

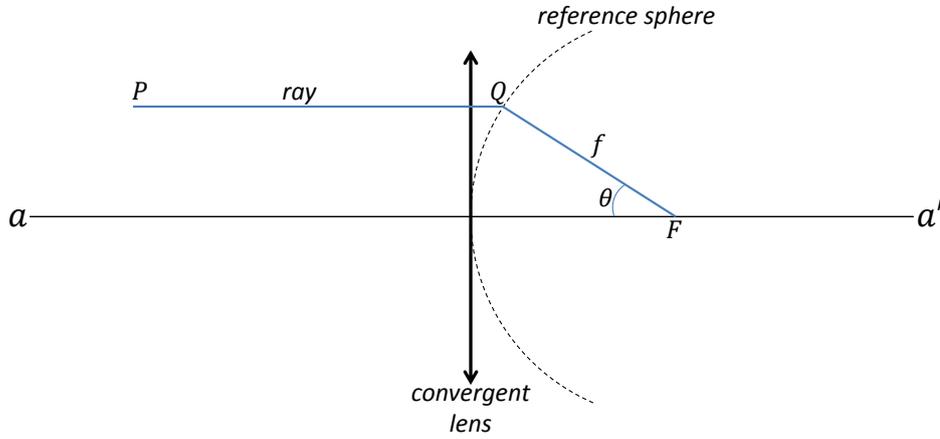


Figure 3.1: Scheme of an aplanatic convergent lens with focal length f and focus at point F . PQF is the path of an optical ray.

PQ and QF are conjugate rays. Moreover, if the media of the half-spaces to the left and right of the lens are denoted with the subscript 1 and 2 respectively, then energy conservation gives:

$$\sqrt{\frac{\epsilon_1}{\mu_1}} \|\vec{E}_1\|^2 dS_1 = \sqrt{\frac{\epsilon_2}{\mu_2}} \|\vec{E}_2\|^2 dS_2$$

where plane wave intensities are assumed for the PQ and QF rays. But:

$$dS_2 = \frac{dS_1}{\cos(\theta)}$$

where $dS_{1,2}$ are element of area centered on the ray paths PQ and QF respectively and perpendicular to them (see Fig. 3.1). Thus the plane wave amplitude of the generic converging optical ray QF , just in correspondence

of the variable point Q on the reference sphere is, as a function of the input field amplitude $\|\vec{E}_1\|$:

$$\|\vec{E}_2\| = \|\vec{E}_1\| \sqrt{\frac{n_1 \mu_2}{n_2 \mu_1} \cos(\theta)}. \quad (3.3)$$

In (3.3) $n_{1,2} = \sqrt{\frac{\epsilon_{1,2} \mu_{1,2}}{\epsilon_0 \mu_0}}$ are the refractive indices of the media and usually, being $\mu_{1,2} \approx 1$, the ratio of the magnetic permeabilities is neglected. Now, assuming the aa' axis coincides with the y axis of a reference frame with origin in the focal point F , it is possible to apply equation (2.8) to calculate the electromagnetic optically focused near field in a region of space around F , using (3.3) as the corresponding far field \vec{E}_∞ . To this end one has to introduce a further angle ϕ , besides θ , to measure the azimuthal rotation around the aa' axis, when the planar integration element $dk_x dk_z$ is replaced with a corresponding, more suitable, differential area element on a spherical surface. If the θ angle is measured starting from the negative y semi-axis (to the left of F , toward the lens) like in Fig. 3.1 — the k_x, k_y, k_z axes coinciding with the x, y, z ones — these planar and spherical elemental areas are related by:

$$dk_x dk_z = \cos(\theta) k^2 \sin(\theta) d\theta d\phi$$

with $\|\vec{k}\| = k$ and the $\frac{1}{\cos(\theta)}$ factor accounting for the projection of the area element on the $k_x k_z$ plane. But, being $k_y = -k \cos(\theta)$, one has:

$$\frac{1}{k_y} dk_x dk_z = -k \sin(\theta) d\theta d\phi$$

with which to replace in the double integration of (2.8). Replacing in it also r with f , one gets:

$$\vec{E}_F(\vec{r}; k) = -\sqrt{\frac{n_1}{n_2}} \frac{ikf e^{-ikf}}{2\pi} \int_0^{\theta_{max}} \int_0^{2\pi} \vec{E}_2(\theta, \phi; k) e^{ik\hat{u}\cdot\vec{r}} \cos^{\frac{1}{2}}(\theta) \sin(\theta) d\theta d\phi$$

where:

- θ_{max} is an aperture upper bound due to the finite size of the lens (aperture stop in a screen and its entrance pupil, which is the image of the entrance stop);
- \hat{u} is a unit vector pointing from F in the various directions of the solid integration angle;
- \vec{E}_2 is the plane wave field vector “at infinity”, i.e. at points on the reference sphere. It lies in a plane perpendicular to QF . Its strength, by (3.3), depends

on that, \vec{E}_1 , of the impinging beam on the lens;

- \vec{r} is the position vector from the focal point F to the point in space in which \vec{E}_F is calculated.

With reference to Fig. 3.2, there is a change from a cylindrical to a spherical geometry at the reference sphere:

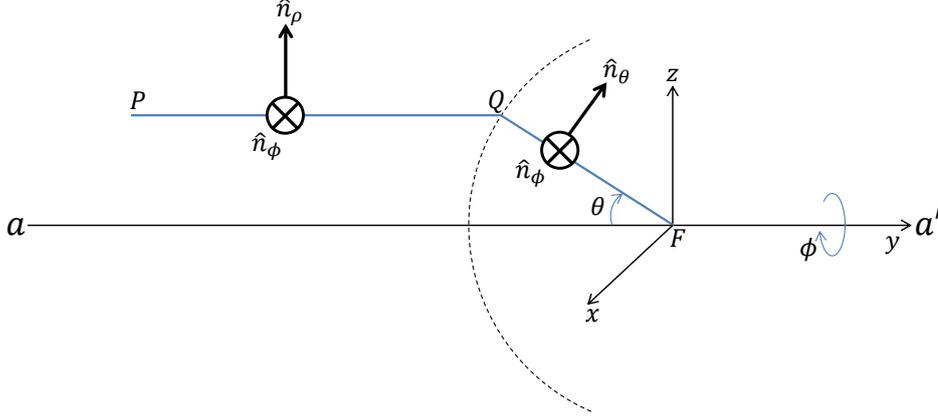


Figure 3.2: Reference frame and unit vectors for the electromagnetic field of a focusing aplanatic lens.

One sees that the unit vector \hat{n}_ϕ is unaffected while \hat{n}_ρ transforms into \hat{n}_θ . If the angle ϕ is measured starting from the x axis in a clockwise sense (i.e., opposite with respect to the right-hand grip rule with the positive y axis as the thumb), using a fixed basis \hat{x} , \hat{y} and \hat{z} one has in general:

$$\hat{u} = \sin(\theta) \cos(\phi) \hat{x} - \cos(\theta) \hat{y} + \sin(\theta) \sin(\phi) \hat{z}$$

$$\hat{n}_\rho = \cos(\phi) \hat{x} + \sin(\phi) \hat{z}$$

$$\hat{n}_\phi = -\sin(\phi) \hat{x} + \cos(\phi) \hat{z}$$

$$\hat{n}_\theta = \cos(\theta) \cos(\phi) \hat{x} + \sin(\theta) \hat{y} + \cos(\theta) \sin(\phi) \hat{z}.$$

Thus it follows that:

$$\vec{E}_2(\theta, \phi; k) = t_{te}(\theta) \left(\vec{E}_1 \cdot \hat{n}_\phi \right) \hat{n}_\phi + t_{tm}(\theta) \left(\vec{E}_1 \cdot \hat{n}_\rho \right) \hat{n}_\theta$$

where t_{te} and t_{tm} are the lens transmission amplitudes for the TE (transverse electric) and TM (transverse magnetic) polarizations respectively (see Chapter 4, Section 4.1). In the present case it will be considered the case of an

ideal lens with $t_{te} = t_{tm} = 1$ and an incident field $\vec{E}_1(\theta, \phi; k)$ linearly polarized along the z direction, with a separated k dependence for the amplitude. As an example, the gaussian monochromatic field distribution of (3.2) with the waist in correspondence of the plane of the lens, which means $y = 0$ in (3.2):

$$\vec{E}_1(\theta, \phi; k) = \hat{z} E_o(k) e^{-\frac{f^2 \sin^2(\theta)}{w_o^2}} = \hat{z} E_o(k) g(\theta, \phi).$$

In fact, x and z in (3.2) are $x = f \sin(\theta) \cos(\phi)$ and $z = f \sin(\theta) \sin(\phi)$ respectively for small aperture angles. With this choice one has:

$$\vec{E}_2(\theta, \phi; k) = E_o(k) g(\theta, \phi) [\cos(\phi) \hat{n}_\phi + \sin(\phi) \hat{n}_\theta] = E_o(k) g(\theta, \phi) \vec{w}(\theta, \phi).$$

By denoting collectively:

$$\vec{G}(\theta, \phi) = g(\theta, \phi) \vec{w}(\theta, \phi)$$

and remembering that $k = \frac{\omega}{c}$ where c is the phase velocity inside the medium to the left of the lens (with subscript 2), the above double integral for the focused optical field near F can be written, assuming for the sake of simplicity $n_1 = n_2 = 1$, as:

$$\vec{E}_F(\vec{r}; \omega) = -\frac{i\omega f}{2\pi c_o} \int_0^{\theta_{max}} \int_0^{2\pi} E_o(\omega) \vec{G}(\theta, \phi) e^{-i\frac{\omega}{c_o}(f - \hat{u} \cdot \vec{r})} \cos^{\frac{1}{2}}(\theta) \sin(\theta) d\theta d\phi$$

where c_o is the speed of light in vacuum. Multiplying the above expression by the time phase factor $e^{-i\omega t}$, integrating over t and remembering the correspondence:

$$i\omega \longleftrightarrow -\frac{\partial}{\partial t},$$

it becomes possible to express the optical field in the time domain (by an abuse of notation, the same letters for the time and frequency domain fields are used) as:

$$\boxed{\vec{E}_F(\vec{r}, t) = \frac{f}{2\pi c_o} \int_0^{\theta_{max}} \int_0^{2\pi} \dot{E}_o \left(t' + \frac{\hat{u} \cdot \vec{r}}{c_o} \right) \vec{G}(\theta, \phi) \cos^{\frac{1}{2}}(\theta) \sin(\theta) d\theta d\phi} \quad (3.4)$$

where the dot above a letter denotes time differentiation (some normalization factors have been reabsorbed in E_o) and $t' = t - \frac{f}{c_o}$ (a time shift). The

importance of (3.4) resides in the fact it represent the focused optical field as a superposition of plane waves propagating in the direction *opposite* to that pointed by \hat{u} as it scans, with its tip, the given solid angle. In fact, for each θ and ϕ fixed, \vec{G} , which gives the direction of the \vec{E} field of the superposing waves, is orthogonal to \hat{u} . Expression (3.4) is similar to those given in [41, 42] (see also [43, 44]). The function $E_o(t)$ defines the time profile of the incident beam in the plane of the lens.

Chapter 4

Planarly multilayered media

4.1 Matrix method for \mathcal{R} and \mathcal{T}

This section is devoted to the characterization of the reflectance \mathcal{R} and the transmittance \mathcal{T} of N contiguous layers stacked along the y axis, each of finite thickness d_s ($s = 1, \dots, N$) and infinite extension in the xz plane, in presence of a beam of light described by a plane electromagnetic wave impinging on them with a given angle of incidence. The various layers are made of a homogeneous material with a constant absolute *complex* permittivity ϵ_s :

$$\epsilon_s = \epsilon'_s + i\epsilon''_s$$

(i is the imaginary unit and the plus sign is coherent with the choice of the sign in the time harmonic phase factor). The complex index of refraction is:

$$n_s = n'_s + in''_s = \sqrt{\frac{\epsilon_s \mu_s}{\epsilon_o \mu_o}} = \sqrt{\frac{\epsilon_s}{\epsilon_o}}.$$

($s = 1, \dots, N$) where the last equality holds for non-magnetic ($\mu \equiv \mu_o$ everywhere, the vacuum permeability) materials only. Also included are a semi-infinite non-absorptive $s = 0$ layer with real index of refraction n_0 , which extends to $y = -\infty$, containing the incident and reflected plane waves, and a semi-infinite $s = N + 1$ layer which extends to $y = +\infty$, containing the transmitted plane wave. We thus have an overall sequence of $N + 2$ media identified by the integer $s = 0, \dots, N + 1$. The xy plane is assumed to be the plane of incidence. Two cases have to be considered: TE-polarization (or s-wave, from “senkrecht” that means perpendicular in German) with the electric field vector \vec{E} along the z -axis, and TM-polarization (or p-wave, from parallel) with the magnetic field vector \vec{H} along the z -axis.

• **Single interface:** with a single interface only, the situation is that depicted in Figs 4.1 and 4.2, where on each side ($s = 1$ or $s = 2$) a general superposition of an ingoing and an outgoing plane wave is considered. The field amplitudes on either sides are:

$$\mathbf{TE:} \quad E_{s,z} = [E_s^{(+)} e^{+ik_s,y y} + E_s^{(-)} e^{-ik_s,y y}] e^{ik_s,x x}$$

$$\mathbf{TM:} \quad H_{s,z} = Z_s^{-1} [-E_s^{(+)} e^{+ik_s,y y} + E_s^{(-)} e^{-ik_s,y y}] e^{ik_s,x x}$$

($s = 1, 2$; the minus sign has been introduced because the vector is entering the plane of the sheet) and Z_s is the medium characteristic impedance:

$$Z_s = \sqrt{\frac{\mu_s}{\epsilon_s}} = \sqrt{\frac{\mu_0}{\epsilon_s}}$$

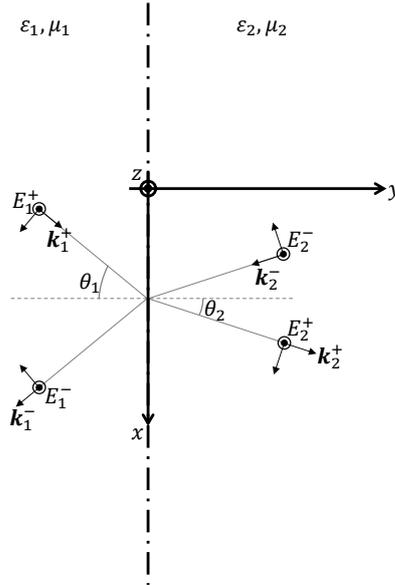


Figure 4.1: TE polarized light on a single interface between two media (\vec{E} points upward from the plane of the sheet, which is the plane of incidence).

where, again, the last equality holds for non-magnetic materials only. By using the curl operator from the Maxwell's equations in the frequency domain one gets for the *tangential* components:

$$\mathbf{TE:} \quad H_{s,x} = \frac{1}{i\omega\mu_s} (\vec{\nabla} \times \vec{E}_s)_x = \frac{k_{s,y}}{\omega\mu_s} [E_s^{(+)} e^{+ik_s,y y} - E_s^{(-)} e^{-ik_s,y y}] e^{ik_s,x x}$$

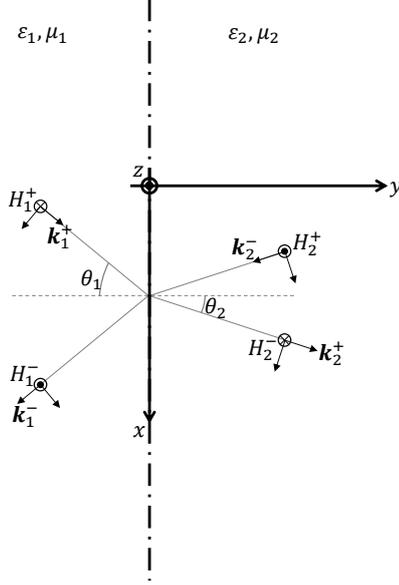


Figure 4.2: TM polarized light on a single interface between two media (\vec{H} points upward from the plane of the sheet, which is the plane of incidence).

$$\mathbf{TM:} \quad E_{s,x} = -\frac{1}{i\omega\epsilon_s}(\vec{\nabla} \times \vec{H}_s)_x = \frac{Z_s^{-1}k_{s,y}}{\omega\epsilon_s} [E_s^{(+)}e^{+ik_{s,y}y} + E_s^{(-)}e^{-ik_{s,y}y}] e^{ik_{s,x}x}.$$

($s = 1, 2$). Here the wave vector lies in the xy plane of incidence:

$$\vec{k}_s = \hat{x}k_{s,x} + \hat{y}k_{s,y} \quad (4.1)$$

and is in general a complex quantity satisfying:

$$\vec{k}_s \cdot \vec{k}_s = \left(\frac{\omega}{c_o}n_s\right)^2 \quad (4.2)$$

where c_o is the vacuum light speed. By applying the continuity condition of the tangential z and x field components for any coordinates at the interface plane $y = 0$, one gets:

$$k_{1,x} = k_{2,x}$$

which is the usual Snell's law: $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$, if both materials are non dissipative. If the $s = 2$ medium is absorptive then $k_{2,x}$, $k_{2,y}$ and the angle θ_2 are complex quantities. The ray interpretation on the $s = 2$ side of Figs 4.1 and 4.2 could be misleading in this case, because the transmitted wave

is non-homogeneous with equal amplitude and equal phase planes having different orientations. For the field amplitudes one instead gets:

$$\mathbf{TE} : \begin{bmatrix} 1 & 1 \\ \frac{k_{1,y}}{\mu_1} & -\frac{k_{1,y}}{\mu_1} \end{bmatrix} \begin{bmatrix} E_1^{(+)} \\ E_1^{(-)} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \frac{k_{2,y}}{\mu_2} & -\frac{k_{2,y}}{\mu_2} \end{bmatrix} \begin{bmatrix} E_2^{(+)} \\ E_2^{(-)} \end{bmatrix} \quad (4.3a)$$

$$\mathbf{TM} : \frac{1}{Z_1} \begin{bmatrix} \frac{k_{1,y}}{\epsilon_1} & \frac{k_{1,y}}{\epsilon_1} \\ 1 & -1 \end{bmatrix} \begin{bmatrix} E_1^{(+)} \\ E_1^{(-)} \end{bmatrix} = \frac{1}{Z_2} \begin{bmatrix} \frac{k_{2,y}}{\epsilon_2} & \frac{k_{2,y}}{\epsilon_2} \\ 1 & -1 \end{bmatrix} \begin{bmatrix} E_2^{(+)} \\ E_2^{(-)} \end{bmatrix} \quad (4.3b)$$

where use has been made of matrix notation. Let us call D_s ($s = 1, 2$) the 2×2 matrices involved in the above linear relations (4.3) (those in (4.3a) for the TE case, those in (4.3b) for the TM case: this is left understood in all what follows); this allows one to obtain a linear relation among $E_1^{(+)}$, $E_1^{(-)}$, $E_2^{(+)}$ and $E_2^{(-)}$:

$$\begin{bmatrix} E_1^{(+)} \\ E_1^{(-)} \end{bmatrix} = (D_1)^{-1} D_2 \begin{bmatrix} E_2^{(+)} \\ E_2^{(-)} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} E_2^{(+)} \\ E_2^{(-)} \end{bmatrix}$$

from which to calculate the reflection and transmission *amplitudes*:

$$r = \frac{E_1^{(-)}}{E_1^{(+)}} = \frac{c}{a}$$

$$t = \frac{E_2^{(+)}}{E_1^{(+)}} = \frac{1}{a}$$

both calculated imposing the condition: $E_2^{(-)} = 0$ (note that here, for TM-polarization, the ratio of the electric field amplitudes has been considered and not that of the magnetic ones. This choice is due to the instrumental response, which depends on the electric field). They would give the so called Fresnel coefficients. The single interface example serves as the building block for a sequence of layers.

• **Single layer:** for $N = 1$ homogeneous layers ($s = 0, 1, 2$), the situation is that depicted in Fig. 4.3, with a near (left) and a far (right) interface of a single slab of thickness $d_1 = L$.

The field phases inside the (each, when $N > 1$) slab is assumed as indicated in Fig. 4.3. By applying at the left and right interfaces of the middle $s = 1$ layer relations like (4.3), but also with the insertion of the matrix P_1 connecting the amplitudes at the two interfaces (the choice of signs is coherent with that in the time harmonic phase factor):

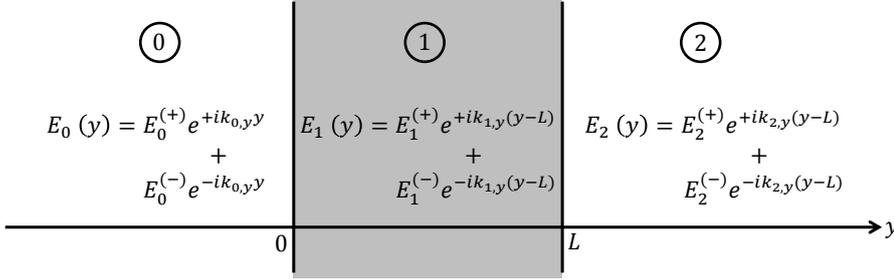


Figure 4.3: Single ($N = 1$) layer with two interfaces. The electric field expressions are for the TE case. For the TM case Z_s factors should appear. The phase factors are the same for both the TE and the TM cases.

$$\begin{bmatrix} E_1^{(+)} \\ E_1^{(-)} \end{bmatrix}_{\text{left}} = P_1 \begin{bmatrix} E_1^{(+)} \\ E_1^{(-)} \end{bmatrix}_{\text{right}} = \begin{bmatrix} e^{-ik_1 y L} & 0 \\ 0 & e^{+ik_1 y L} \end{bmatrix} \begin{bmatrix} E_1^{(+)} \\ E_1^{(-)} \end{bmatrix}_{\text{right}},$$

one gets after matrix inversion:

$$\begin{bmatrix} E_0^{(+)} \\ E_0^{(-)} \end{bmatrix} = (D_0)^{-1} \{ D_1 P_1 (D_1)^{-1} \} D_2 \begin{bmatrix} E_2^{(+)} \\ E_2^{(-)} \end{bmatrix}.$$

The use of the propagation matrix P_1 amounts to assume the right interfaces as the phase reference planes inside each layer (finite or semi-infinite), with the exception of the last one, in which the left interface plane (obviously) is the phase reference (see again the field expressions in Fig. 4.3).

• **Multilayer:** from the previous example a pattern emerges, because the above relation can be in a straightforward manner generalized to N layers, by iterating N times the product of the matrices in the curly braces:

$$\begin{bmatrix} E_0^{(+)} \\ E_0^{(-)} \end{bmatrix} = (D_0)^{-1} \Delta D_{N+1} \begin{bmatrix} E_{N+1}^{(+)} \\ E_{N+1}^{(-)} \end{bmatrix}$$

where:

$$\Delta = \prod_{k=1}^N \{ D_k P_k (D_k)^{-1} \}.$$

The calculation of the overall matrix S for a stack of N layers plus the initial and final media:

$$S = (D_0)^{-1} \Delta D_{N+1} \quad (4.4)$$

is well suitable for implementation on a computer after specification of:

- 1) the electric and magnetic (possibly complex) parameters ϵ_s, μ_s of the various N layers at a given angular frequency ω ;
- 2) their thicknesses d_s ;
- 3) the angle of incidence θ_0 of the incoming plane wave in medium $s = 0$, being assumed to be non-absorptive, which allows, through the chain of equalities (the generalization of the one already seen for the $N = 1$ case):

$$k_{s,x} = k_{s+1,x} \quad (s = 1, \dots, N), \quad (4.5)$$

to calculate the (complex) angle θ_s in each layer and thus the elements of D_s and its inverse. In fact from (4.1) and (4.2) one has:

$$k_{s,x} = \frac{\omega}{c_o} (n'_s + in''_s) \sin(\theta_s)$$

$$k_{s,y} = \frac{\omega}{c_o} (n'_s + in''_s) \cos(\theta_s)$$

and by (4.5), for each $s = 1, \dots, N + 1$:

$$\sin(\theta_s) = \frac{n'_0}{n'_s + in''_s} \sin(\theta_0)$$

$$\cos(\theta_s) = \sqrt{1 - \sin^2(\theta_s)}.$$

From these it follows that:

$$\cos(\theta_s) = \sqrt{1 - \frac{n_0'^2(n_s'^2 - n_s''^2)}{(n_s'^2 + n_s''^2)^2} \sin^2(\theta_0) + i \frac{2n_s' n_s'' n_0'^2}{(n_s'^2 + n_s''^2)^2} \sin^2(\theta_0)}.$$

By putting also:

$$\cos(\theta_s) = q_s e^{i\gamma_s}$$

q_s and γ_s real with $q_s > 0$, one has:

$$\cos^2(\theta_s) = q_s^2 e^{2i\gamma_s} = q_s^2 \cos(2\gamma_s) + iq_s^2 \sin(2\gamma_s)$$

and thus by squaring the previous expression for $\cos(\theta_s)$ and equating with the above expression one gets:

$$q_s^2 \cos(2\gamma_s) = 1 - \frac{n_0'^2(n_s'^2 - n_s''^2)}{(n_s'^2 + n_s''^2)^2} \sin^2(\theta_0)$$

$$q_s^2 \sin(2\gamma_s) = \frac{2n'_s n''_s n_0^2}{(n_s'^2 + n_s''^2)^2} \sin^2(\theta_0)$$

from which q_s and γ_s can be calculated as functions of θ_0 , n'_0 , n'_s and n''_s . From these expressions, those for $U_s \equiv \text{Re} \{(n'_s + in''_s) \cos(\theta_s)\} = \frac{c_0}{\omega} \text{Re} \{k_{s,y}\}$ and $V_s \equiv \text{Im} \{k_{s,y}\}$ can be obtained:

$$U_s = n'_s q_s \cos(\gamma_s) - n''_s q_s \sin(\gamma_s)$$

$$V_s = \frac{\omega}{c_0} [n'_s q_s \sin(\gamma_s) + n''_s q_s \cos(\gamma_s)] .$$

The reflectance \mathcal{R} and transmittance \mathcal{T} can be calculated by means of the RMS power flux \mathcal{P}_u , along the u axis, from the real part of the complex Poynting vector:

$$\mathcal{P}_u = \frac{1}{2} \text{Re} \left\{ \left(\vec{E} \times \vec{H}^* \right) \cdot \hat{u} \right\}$$

where $u = x, y, z$. If $d_{tot} = \sum_{s=1}^N d_s$ is the total thickness of the multilyer, which starts at $y = 0$, and

$$\vec{k}'_s = \hat{x} k_{s,x} - \hat{y} k_{s,y} ,$$

the fields are (for non-magnetic materials):

TE-polarization)

$$\vec{E}_{inc} = \hat{z} E_0^{(+)} e^{i(k_{0,x}x + k_{0,y}y)}$$

$$\vec{H}_{inc} = \frac{1}{\omega \mu_0} \vec{k}_0 \times \vec{E}_{inc}$$

$$\vec{E}_{refl} = \hat{z} r E_0^{(+)} e^{i(k_{0,x}x - k_{0,y}y)}$$

$$\vec{H}_{refl} = \frac{1}{\omega \mu_0} \vec{k}'_0 \times \vec{E}_{refl}$$

$$\vec{E}_{tr} = \hat{z} t E_0^{(+)} e^{i[k_{N+1,x}x + k_{N+1,y}(y - d_{tot})]}$$

$$\vec{H}_{tr} = \frac{1}{\omega \mu_0} \vec{k}_0 \times \vec{E}_{tr}$$

TM-polarization)

$$\begin{aligned}
\vec{E}_{inc} &= -\frac{1}{\omega\epsilon_0}\vec{k}_0 \times \vec{H}_{inc} \\
\vec{H}_{inc} &= -\hat{z}\frac{1}{Z_0}E_0^{(+)}e^{i(k_{0,x}x+k_{0,y}y)} \\
\vec{E}_{refl} &= -\frac{1}{\omega\epsilon_0}\vec{k}'_0 \times \vec{H}_{refl} \\
\vec{H}_{refl} &= \hat{z}\frac{1}{Z_0}rE_0^{(+)}e^{i(k_{0,x}x-k_{0,y}y)} \\
\vec{E}_{tr} &= -\frac{1}{\omega\epsilon_{N+1}}\vec{k}_{N+1} \times \vec{H}_{tr} \\
\vec{H}_{tr} &= -\hat{z}\frac{1}{Z_{N+1}}tE_0^{(+)}e^{i[k_{N+1,x}x+k_{N+1,y}(y-d_{tot})]}.
\end{aligned}$$

Taking account that, by (4.5), $k_{0,x}$ is real (the $s = 0$ medium is non-absorptive), that $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$, and by defining ξ as the distance ($\xi \geq 0$) from the interface at $y = d_{tot}$, one can get the following power fluxes at $y = 0$ (incident and reflected) and at $y = d_{tot}$ (transmitted) for non-magnetic materials:

$$\begin{aligned}
\mathcal{P}_{inc,x} &= \frac{|E_0^{(+)}|^2}{2c_o\mu_o}n'_0 \sin(\theta_0) \\
\mathcal{P}_{inc,y} &= \frac{|E_0^{(+)}|^2}{2c_o\mu_o}n'_0 \cos(\theta_0) \\
\mathcal{P}_{refl,x} &= \frac{|r|^2|E_0^{(+)}|^2}{2c_o\mu_o}n'_0 \sin(\theta_0) \\
\mathcal{P}_{refl,y} &= -\frac{|r|^2|E_0^{(+)}|^2}{2c_o\mu_o}n'_0 \cos(\theta_0) \\
\mathcal{P}_{tr,x} &= \frac{|t|^2|E_0^{(+)}|^2}{2c_o\mu_o}n'_0 \sin(\theta_0)e^{-2Im\{k_{N+1,y}\}\xi} \\
\mathcal{P}_{tr,y} &= \frac{|t|^2|E_0^{(+)}|^2}{2c_o\mu_o}\frac{c_o}{\omega}Re\{k_{N+1,y}\}e^{-2Im\{k_{N+1,y}\}\xi}.
\end{aligned}$$

The multilayer reflectance \mathcal{R} and transmittance \mathcal{T} are thus given by:

$$\mathcal{R} = |r|^2$$

$$\mathcal{T} = |t|^2 \frac{U_{N+1}}{n'_0 \cos(\theta_0)} e^{-2V_{N+1}\xi}.$$

As can be seen, if $\xi \neq 0$ and the final medium $s = N + 1$ is dissipative or in the multilayer ($s = 1, \dots, N$) there is some dissipation and the incidence is not normal, there will be some extra absorption. If $\theta_0 = 0$ (normal incidence) there is no difference between the TE and TM polarization. In any case the overall matrix (4.4) has to be calculated:

$$S = \begin{bmatrix} s_{1,1} & s_{1,2} \\ s_{2,1} & s_{2,2} \end{bmatrix}$$

and the reflection and transmission complex amplitudes evaluated:

$$r = \frac{s_{2,1}}{s_{1,1}}$$

$$t = \frac{1}{s_{1,1}}.$$

4.2 \vec{E} and \vec{H} field distributions

Once the reflection and transmission amplitudes r , t have been calculated, the fields inside each layer can recursively be found by applying N times the relations (4.3a) or (4.3b), implying the D_s matrices, starting for example from the left with:

$$E_0^{(+)} \text{ arbitrary}$$

$$E_0^{(-)} = rE_0^{(+)}$$

which correspond to a field in the initial, $s = 0$, semi-infinite medium given by:

$$\mathbf{TE:} \quad E_{0,z} = E_0^{(+)} [e^{+ik_{0,y}y} + re^{-ik_{0,y}y}] e^{ik_{0,x}x}$$

$$\mathbf{TM:} \quad H_{0,z} = Z_0^{-1} E_0^{(+)} [-e^{+ik_{0,y}y} + re^{-ik_{0,y}y}] e^{ik_{0,x}x}$$

which hold for:

$$y \leq 0.$$

Proceeding recursively one has for the field inside the s -th layer, in terms of the amplitudes of the $(s - 1)$ -th medium:

$$\begin{bmatrix} E_s^{(+)} \\ E_s^{(-)} \end{bmatrix} = (D_s P_s)^{-1} D_{s-1} \begin{bmatrix} E_{s-1}^{(+)} \\ E_{s-1}^{(-)} \end{bmatrix}$$

which correspond to a field:

$$\mathbf{TE:} \quad E_{s,z} = [E_s^{(+)} e^{+ik_{s,y}(y-y_s)} + E_s^{(-)} e^{-ik_{s,y}(y-y_s)}] e^{ik_{s,x}x}$$

$$\mathbf{TM:} \quad H_{s,z} = Z_s^{-1} [-E_s^{(+)} e^{+ik_{s,y}(y-y_s)} + E_s^{(-)} e^{-ik_{s,y}(y-y_s)}] e^{ik_{s,x}x}$$

where:

$$y_s = \sum_{\ell=1}^s d_\ell$$

with $y_N = d_{tot}$, the total thickness and:

$$y_{s-1} \leq y \leq y_s.$$

In the final, $s = N + 1$, semi-infinite medium the field is:

$$\mathbf{TE:} \quad E_{N+1,z} = tE_0^{(+)} e^{+ik_{s,y}(y-d_{tot})} e^{ik_{N+1,x}x}$$

$$\mathbf{TM:} \quad H_{N+1,z} = -Z_{N+1}^{-1} tE_0^{(+)} e^{+ik_{s,y}(y-d_{tot})} e^{ik_{N+1,x}x}$$

with

$$y \geq d_{tot}.$$

The missing fields can be calculated through $(\vec{\nabla} \times \hat{z}E_{s,z})/i\omega\mu_s$ for the TE case and $-(\vec{\nabla} \times \hat{z}H_{s,z})/i\omega\epsilon_s$ for the TM case, $s = 0, \dots, N + 1$.

4.3 The inverse transmittance problem

As a proposal, the implemented matrix algorithm previously described for evaluating \mathcal{R} , \mathcal{T} of a multilayer system, could be applied for the inverse problem of determining the refractive index of an unknown material on a given region of the visible spectrum, from measured values of the transmittance \mathcal{T} in that range of wavelengths for a slab fabricated with that material. For sake of definiteness, a trilayer system is considered like the one depicted in Fig. 4.4:

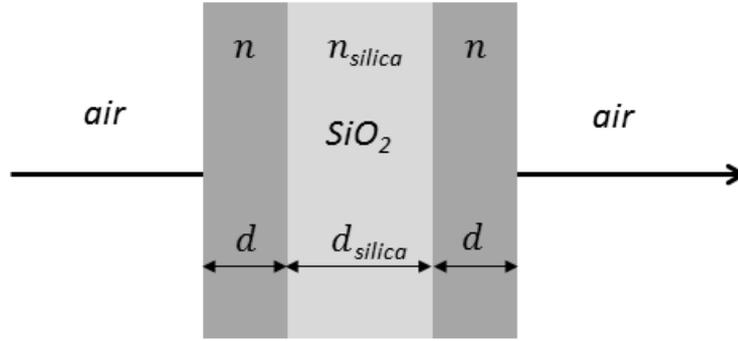


Figure 4.4: Schematic of the measurement....

on which a light beam impinges and from which the experimental transmittance is reported. It is a trilayer . Sellmeier [] formula:

$$n(\lambda) = \sqrt{1 + \frac{H_1\lambda^2}{\lambda^2 + K_1^2} + \frac{H_2\lambda^2}{\lambda^2 + K_2^2} + \frac{H_3\lambda^2}{\lambda^2 + K_3^2}}$$

One thus has the analytical transmittance of the trilayer as a scalar-valued function:

$$\mathcal{T} = \mathcal{T}(d, H_1, K_1, H_2, K_2, H_3, K_3|\lambda)$$

depending nonlinearly from 7 variables and containing also 8 more parameters which are the wavelength λ (explicitly indicated in the expression above), the seven parameters H_i^{silica} , K_i^{silica} ($i = 1, 2, 3$) of the Sellmeier formula for the silica and the thickness of the silica substrate d_{silica} , which are fixed. By considering M wavelength values λ_r ($r = 1, \dots, M$) on a given range $\lambda_{min} \leq \lambda \leq \lambda_{max}$, ordered in increasing values and maybe evenly distributed:

$$\lambda_r = \lambda_{min} + (r - 1) \left(\frac{\lambda_{max} - \lambda_{min}}{M - 1} \right)$$

one can then consider the vector-valued function (written as a column vector):

$$\begin{bmatrix} \mathcal{T}(d, H_1, K_1, H_2, K_2, H_3, K_3 | \lambda_1) \\ \vdots \\ \mathcal{T}(d, H_1, K_1, H_2, K_2, H_3, K_3 | \lambda_M) \end{bmatrix},$$

where every element is the \mathcal{T} function calculated at a given λ_r . By making M transmittance measures \mathcal{T}_r^{meas} at the corresponding wavelengths, one can construct the vector-valued function:

$$\vec{F}(\vec{p}) = \begin{bmatrix} f_1(\vec{p}) \\ \vdots \\ f_M(\vec{p}) \end{bmatrix} = \begin{bmatrix} \mathcal{T}(d, H_1, K_1, H_2, K_2, H_3, K_3 | \lambda_1) - \mathcal{T}_1^{meas} \\ \vdots \\ \mathcal{T}(d, H_1, K_1, H_2, K_2, H_3, K_3 | \lambda_M) - \mathcal{T}_M^{meas} \end{bmatrix},$$

where the vector variable $\vec{p} \in \mathbb{R}^7$ indicates collectively the variables:

$$d, H_1, K_1, H_2, K_2, H_3, K_3.$$

These latter ones will be determined by minimization of the so called *residual norm*, i.e. the scalar valued function:

$$\phi(\vec{p}) = \sum_{r=1}^M f_r^2(\vec{p}) \equiv \|\vec{F}(\vec{p})\|_2^2,$$

which is a nonlinear least squares problem [1], whose solution is the inverse transmittance problem solution. For \vec{p}^* to be a minimum of $\phi(\vec{p})$ it must necessarily be: $\vec{\nabla}\phi(\vec{p}^*) = \vec{0}$, where $\vec{\nabla}\phi(\vec{p})$ is the gradient of $\phi(\vec{p})$. To find \vec{p}^* starting from a suitable near point one can use a second order Taylor approximation around a generic point \vec{p}^o :

$$\phi(\vec{p}^o + \Delta\vec{p}) \approx \vec{\nabla}\phi(\vec{p}^o) + \bar{\bar{H}}(\phi(\vec{p}^o))\Delta\vec{p}$$

where $\bar{\bar{H}}(\phi(\vec{p}^o))$ is the Hessian 7×7 matrix of $\phi(\vec{p})$:

$$\bar{\bar{H}}(\phi(\vec{p})) = \left[\frac{\partial^2 \phi(\vec{p})}{\partial p_i \partial p_j} \right]$$

$i, j = 1, \dots, 7$ at \vec{p}^o . By requiring the left hand member in the above expansion to be $\vec{0}$, one gets the linear system:

$$\bar{\bar{H}}(\phi(\vec{p}^o))\Delta\vec{p} = -\vec{\nabla}\phi(\vec{p}^o) \quad (4.6)$$

to be solved for the correction $\Delta\vec{p}$ which, when calculated iteratively, should bring an initial guess \vec{p}^o toward \vec{p}^* (Newton method for minimizing $\phi(\vec{p})$, [1]). The gradient of $\phi(\vec{p})$, as a column 7×1 vector, can be written as:

$$\vec{\nabla}\phi(\vec{p}) = 2\bar{\bar{J}}^T(\vec{p})\vec{F}(\vec{p})$$

where $\bar{\bar{J}}(\vec{p})$ is the Jacobian $M \times 7$ matrix of $\vec{F}(\vec{p})$:

$$\bar{\bar{J}}(\vec{p}) = \left[\frac{\partial f_r(\vec{p})}{\partial p_s} \right],$$

with $r = 1, \dots, M$ row index and $s = 1, \dots, 7$ column index, and where T means matrix transposition. The Hessian matrix can instead be written as:

$$\bar{\bar{H}}(\phi(\vec{p})) = 2\bar{\bar{J}}^T(\vec{p})\bar{\bar{J}}(\vec{p}) + 2 \sum_{r=1}^M f_r(\vec{p})\bar{\bar{H}}(f_r(\vec{p})).$$

In the Gauss-Newton method [1] the term containing the sum is ignored because if the point \vec{p} is approaching the optimal point \vec{p}^* the $f_r(\vec{p})$ will be small. Thus (4.6) becomes:

$$\boxed{\bar{\bar{J}}^T(\vec{p})\bar{\bar{J}}(\vec{p})\Delta\vec{p} = -\bar{\bar{J}}^T(\vec{p})\vec{F}(\vec{p})} \quad (4.7)$$

which provides a formula for solving for successive update steps $\Delta\vec{p}$ starting from an initial guess \vec{p}^o . The iteration are stopped when two consecutive corrections differ in the $\|\cdot\|_2$ Euclidean norm of \mathbb{R}^7 by a small quantity. The 7×7 coefficient matrix in (4.7) is symmetric and positive semidefinite. Some methods, like the Levenberg-Marquardt method, use a positive parameter to make that matrix positive definite thus allowing the use of suitable factorization methods to efficiently solve the system for the model update steps $\Delta\vec{p}$ [1]. It should be noted that, as in the present case, due to the complexity of the component functions $f_r(\vec{p})$ of $\vec{F}(\vec{p})$, the derivatives in the Jacobian matrix are not calculated analytically but numerically by finite differencing.

4.4 Continuously varying refractive index

The \mathcal{R} , \mathcal{T} spectral behavior of a medium with a refractive index profile $n(y)$, varying unidimensionally as a function of y (although not strictly as a continuous function), can be analyzed with arbitrary accuracy by means of a stacked sequence of sufficiently thin layers, in such a way of approximating the $n(y)$ function by a piecewise constant function.

Part II

Implementation of the FDTD numerical method

Chapter 5

Discretizing the Maxwell's equations

The Finite-Difference Time-Domain method (or simply FDTD method) determines numerically the solution starting directly from the Maxwell's equations in the time domain (1.1), from which the method's name. More precisely, only the curl equations (1.1c) and (1.1c) are considered, because they govern the evolutive dynamics of the electromagnetic field, the remaining divergence equations (1.1a) and (1.1b) merely representing a constraint for the solution. In fact, it can be shown that, by applying the $\vec{\nabla} \cdot$ operator to both members of the curl equations, the divergence equations are obeyed at any successive instant of time if they so are at the initial time:

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{E}) = -\vec{\nabla} \cdot \frac{\partial \vec{B}}{\partial t} \Rightarrow \vec{\nabla} \cdot \vec{B} = \text{const.}$$

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{H}) = \vec{\nabla} \cdot \frac{\partial \vec{D}}{\partial t} + \vec{\nabla} \cdot \vec{j} \Rightarrow \vec{\nabla} \cdot \vec{D} - \rho = \text{const.}$$

because the divergence of the curl of a vector field vanishes and the space and time derivatives have been exchanged. Moreover, in the second, the validity of the continuity equation (local charge conservation) (1.2) has been assumed. The initial condition in the application of the algorithm is one with $\vec{E} \equiv \vec{0}$, $\vec{H} \equiv \vec{0}$ and $\rho \equiv 0$, for which the above expressions hold true with both constants equal to zero. The two curl equations constitute an hyperbolic linear system of evolutive first order partial differential equations:

$$\frac{\partial}{\partial t} \begin{bmatrix} \vec{E} \\ \vec{H} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_x \frac{\partial}{\partial x} + \mathbf{A}_y \frac{\partial}{\partial y} + \mathbf{A}_z \frac{\partial}{\partial z} + \mathbf{B} \end{bmatrix} \begin{bmatrix} \vec{E} \\ \vec{H} \end{bmatrix} + \begin{bmatrix} \vec{j}/\epsilon \\ \vec{0} \end{bmatrix}$$

for the six dimensional unknown vector of the electromagnetic field components (here written as a column vector), where:

$$\mathbf{A}_\alpha = \begin{bmatrix} \mathbf{O} & \mathbf{G}_\alpha/\epsilon \\ \mathbf{G}_\alpha/\mu & \mathbf{O} \end{bmatrix}$$

$\alpha = x, y, z$ are 6×6 block matrices with:

$$\mathbf{G}_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{G}_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{G}_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

3×3 matrices, with \mathbf{O} the 3×3 null matrix and

$$\mathbf{B} = \begin{bmatrix} \frac{\sigma}{\epsilon} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

with \mathbf{I} the 3×3 identity matrix. The coefficient matrices \mathbf{A}_α , \mathbf{B} depend on the position, being typically piecewise constant function of \vec{r} . Technically the above linear PDEs system is hyperbolic because [19] the matrix:

$$P = P(\xi_x, \xi_y, \xi_z) = \xi_x A_x + \xi_y A_y + \xi_z A_z,$$

with $\xi_\alpha \in \mathbb{R}$ ($\alpha = x, y, z$), is *diagonalizable* at each point (\vec{r}, t) of the space-time domain and the eigenvalues are *real* with the norms of the diagonalizing matrix Ω and Ω^{-1} uniformly bounded with respect to ξ_x, ξ_y, ξ_z (strictly hyperbolic if all the eigenvalues are distinct). From the mathematical physics viewpoint, it constitutes an Initial Boundary Value Problem (IBVP) whose solution can generally be found, as a matter of fact, only numerically.

5.1 Time-domain bulk algorithm

The time-domain bulk algorithm concerns the above hyperbolic system posed in the whole of \mathbb{R}^3 , i.e., without any particular boundary condition imposed on the solution, except that the field components go sufficiently fast to zero when $r \rightarrow \infty$ (i.e., a pure Initial Value Problem). Whichever

the electromagnetic problem at hand, to become numerically tractable it needs to be formulated using discrete degrees of freedom for the governing equations. To this end, after considering a time interval $[0, T]$, the space-time slice $\mathbb{R}^3 \times [0, T]$ is sampled at integer multiples of the steps δ_x , δ_y , δ_z and δ_t . The solution will be calculated only at this sampling points. The actual distribution of these points follows that of Yee [14] and, for what concerns the space \mathbb{R}^3 , can most effectively be thought as obtained from a unit cell (not necessarily cubic) like the one depicted in Fig. 5.1:

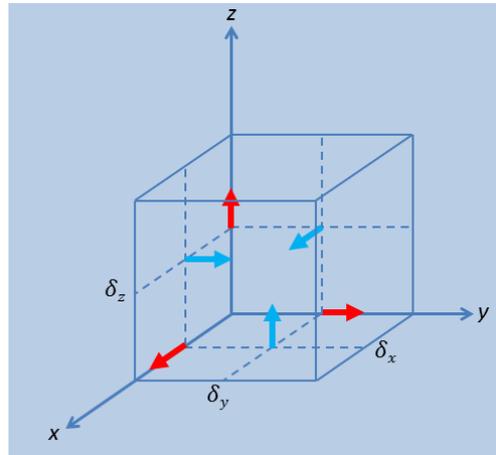


Figure 5.1: A cubic Yee unit cell for \vec{E} (red) and \vec{H} (blue) components sampling in \mathbb{R}^3 .

by iterating it along the coordinate axes. In Fig. 5.1 the red arrows stand for the \vec{E} components and the blue arrows stand for the \vec{H} ones: there is a suitable half space-step shifting among the various components. To get an image of what happens with contiguous Yee cells one can refer to Fig. 5.2.

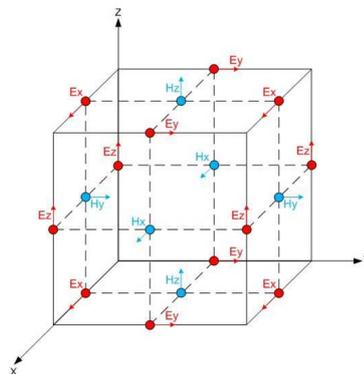


Figure 5.2: A cubic Yee cell with his three far faces made by the contiguous successive cells along the x , y and z axes.

The idea behind such a spatial distribution of sampling points is that to be a “curl conformal” one, with each component of the \vec{E} or \vec{H} vector discretely surrounded by the components of the dual field (respectively, \vec{H} or \vec{E}), circulating around it in a perpendicular plane, like in Fig. 5.3:

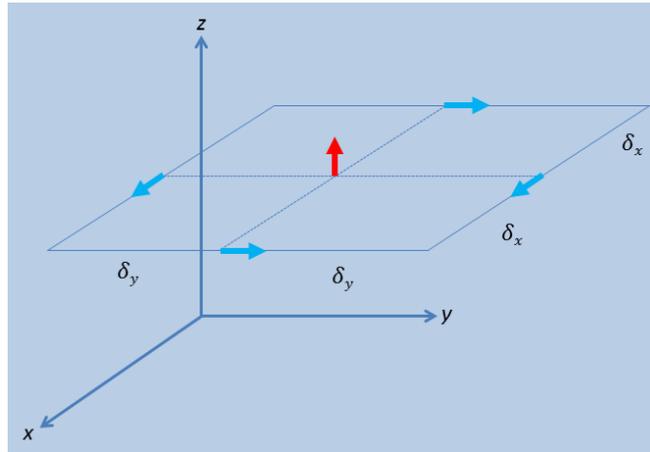


Figure 5.3: The z component of \vec{E} surrounded by four circulating x and y components of \vec{H} .

This distribution allows to write finite difference expressions approximating the partial derivatives in the curl equations (the first order PDEs system above) with second order of accuracy. To get second order of accuracy also in the finite difference expression for the time derivative, a similar trick of half time-step shifting is used for the sampling points along the time t -axis for \vec{E} and \vec{H} , in accordance with Fig. 5.4:

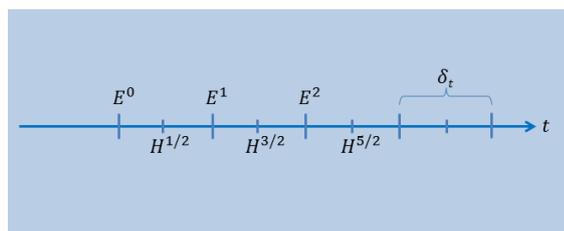


Figure 5.4: Scheme of the time axis \vec{E} and \vec{H} sampling point distribution. Superscripts denote time instants in δ_t units.

For a less cumbersome notation it is convenient — after having identified each Yee cell with the triple (i, j, k) expressing the position of its vertex in $\delta_x, \delta_y, \delta_z$ units in a given reference frame $Oxyz$ — to refer each component of the electric or magnetic fields to the cell to which they belong, leaving implicit the actual position inside it:

- halfway the corresponding edges for \vec{E} ,
 - in the middle of the perpendicular faces for \vec{H}
- (both the three edges and the three faces have in common that cell vertex).
The space dependence of the electric field components by means of cell indices thus becomes:

$$E_x((i + 1/2)\delta_x, j\delta_y, k\delta_z, n\delta_t) \implies E_x^n(i, j, k)$$

$$E_y(i\delta_x, (j + 1/2)\delta_y, k\delta_z, n\delta_t) \implies E_y^n(i, j, k)$$

$$E_z(i\delta_x, j\delta_y, (k + 1/2)\delta_z, n\delta_t) \implies E_z^n(i, j, k),$$

and similarly for the magnetic field components:

$$H_x(i\delta_x, (j + 1/2)\delta_y, (k + 1/2)\delta_z, (n + 1/2)\delta_t) \implies H_x^{n+\frac{1}{2}}(i, j, k)$$

$$H_y((i + 1/2)\delta_x, j\delta_y, (k + 1/2)\delta_z, (n + 1/2)\delta_t) \implies H_y^{n+\frac{1}{2}}(i, j, k)$$

$$H_z((i + 1/2)\delta_x, (j + 1/2)\delta_y, k\delta_z, (n + 1/2)\delta_t) \implies H_z^{n+\frac{1}{2}}(i, j, k).$$

This choice makes easier also the actual implementation of the algorithm code. The time dependence instead, is denoted with a superscript. Henceforth, without lack of generality, it is assumed that only non-magnetic materials are considered, which means $\mu = \mu_o$ everywhere.

The first equation of the PDEs system corresponding to the Maxwell's curl equations is:

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon} \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right) - \frac{\sigma}{\epsilon} E_x - \frac{1}{\epsilon} j_x$$

and by writing finite-differences for the derivatives using the space-time sampling points just described it becomes:

$$\frac{E_x^{n+1}(i, j, k) - E_x^n(i, j, k)}{\delta_t} = \dots$$

$$\dots \frac{1}{\epsilon_x(i, j, k)} \left[\frac{H_z^{n+\frac{1}{2}}(i, j, k) - H_z^{n+\frac{1}{2}}(i, j - 1, k)}{\delta_y} - \dots \right]$$

$$\dots \frac{H_y^{n+\frac{1}{2}}(i, j, k) - H_y^{n+\frac{1}{2}}(i, j, k-1)}{\delta_z} \Big] - \dots$$

$$\dots \frac{\sigma_x(i, j, k)}{\epsilon_x(i, j, k)} \left[\frac{E_x^{n+1}(i, j, k) + E_x^n(i, j, k)}{2} \right] - \frac{1}{\epsilon_x(i, j, k)} j_x^{n+\frac{1}{2}}(i, j, k).$$

About this finite-difference expression one should note:

- Taking into account the effective spatial positioning of the \vec{E} and \vec{H} components inside the Yee cell to which the indices i, j, k refer, there is a coherent *spatial centering* among the various terms in the equation (thanks to the suitable half space-steps shifts), even because the current density \vec{j} components are assumed to be sampled in the same locations as the \vec{E} components. Because \vec{j} represents an impressed current density, its spatial distribution and time behavior are known in advance.
- There is a coherent *time centering* among the various terms, thanks to the half time-step shift between \vec{E} and \vec{H} , and if \vec{j} is temporally sampled like the magnetic field. Also, to coherently center the term corresponding to the ohmic current $\sigma\vec{E}$, an arithmetic mean between \vec{E}^n and \vec{E}^{n+1} has been introduced.
- ϵ and σ appear with a component subscript, but that does not mean they are tensors. It means the points of the space grid are assigned to the various media using a *per component* criterion. It could well be that, due to its geometric characteristics and the $\delta_x, \delta_y, \delta_z$ sizes, the boundary between two different media is such that the actual sampling points for \vec{E} in a Yee cell fall on different sides. In any case, it is clear that the material structure — the variability of the coefficients in the first order PDEs system — is realized by means of the discrete variability of ϵ and σ on the space grid, according to an approximation of given continuous distributions in the model.
- By applying Taylor series expansion for multivariate functions of x, y, z and t variables, one can easily see that all terms in both members of the finite-difference equation are approximations of the corresponding ones in the analytic continuous equation, the residual terms neglected being $o(\delta_y^2)$ or $o(\delta_z^2)$ or $o(\delta_t^2)$. That is, by appropriately staggering in space and time the sampling point for the components of the electric and magnetic fields, a second order accuracy is achieved.

The third equation of the PDEs system is:

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu_o} \left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} \right)$$

and by writing finite-differences for the derivatives using the space-time sampling points just described it becomes:

$$\frac{H_x^{n+\frac{1}{2}}(i, j, k) - H_x^{n-\frac{1}{2}}(i, j, k)}{\delta_t} = \frac{1}{\mu_o} \left[\frac{E_y^n(i, j, k+1) - E_y^n(i, j, k)}{\delta_z} - \dots \right. \\ \left. \dots \frac{E_z^n(i, j+1, k) - E_z^n(i, j, k)}{\delta_y} \right],$$

to which the same observations as in the preceding case, *mutatis mutandis*, applies. In particular, second order accuracy — residual terms are $o(\delta_y^2)$ or $o(\delta_z^2)$ or $o(\delta_t^2)$ — continues to apply. As above, spatial indices i, j, k refer to Yee cells.

A similar pattern can be repeated for the remaining equations for \vec{E} and \vec{H} components, with the same characteristics. By resolving the difference equations for $\vec{H}^{n+\frac{1}{2}}$ and \vec{E}^{n+1} , one gets a consistent, second order accurate, *explicit scheme* by means of which to update, at each space location, first the preexisting \vec{H} values, then the preexisting \vec{E} ones in an iterative, time marching way, thus tracking the time evolution of the electromagnetic fields, starting from an initial configuration (which is usually an $\vec{H} = \vec{E} = \vec{0}$ configuration). The excitation of the system is “injected” by means of the current density \vec{j} or by generating an electromagnetic plane wave impinging on a target by means of the so called total field/scattered field method, which will be described in a next Chapter.

Below are reported the FDTD updating equations for the field components of the magnetic field:

$$H_x^{n+\frac{1}{2}}(i, j, k) = H_x^{n-\frac{1}{2}}(i, j, k) + \frac{\delta_t}{\mu_o} \left[\frac{E_y^n(i, j, k+1) - E_y^n(i, j, k)}{\delta_z} - \frac{E_z^n(i, j+1, k) - E_z^n(i, j, k)}{\delta_y} \right] \quad (5.1a)$$

$$H_y^{n+\frac{1}{2}}(i, j, k) = H_y^{n-\frac{1}{2}}(i, j, k) + \frac{\delta_t}{\mu_o} \left[\frac{E_z^n(i+1, j, k) - E_z^n(i, j, k)}{\delta_x} - \frac{E_x^n(i, j, k+1) - E_x^n(i, j, k)}{\delta_z} \right] \quad (5.1b)$$

$$H_z^{n+\frac{1}{2}}(i, j, k) = H_z^{n-\frac{1}{2}}(i, j, k) + \frac{\delta t}{\mu_o} \left[\frac{E_x^n(i, j+1, k) - E_x^n(i, j, k)}{\delta_y} - \frac{E_y^n(i+1, j, k) - E_y^n(i, j, k)}{\delta_x} \right] \quad (5.1c)$$

and of the electric field:

$$E_x^{n+1}(i, j, k) = P_x(i, j, k)E_x^n(i, j, k) + Q_x(i, j, k) \left[\frac{H_z^{n+\frac{1}{2}}(i, j, k) - E_z^{n+\frac{1}{2}}(i, j-1, k)}{\delta_y} - \frac{H_y^{n+\frac{1}{2}}(i, j, k) - H_y^{n+\frac{1}{2}}(i, j, k-1)}{\delta_z} \right] \quad (5.2a)$$

$$E_y^{n+1}(i, j, k) = P_y(i, j, k)E_y^n(i, j, k) + Q_y(i, j, k) \left[\frac{H_x^{n+\frac{1}{2}}(i, j, k) - E_x^{n+\frac{1}{2}}(i, j, k-1)}{\delta_z} - \frac{H_z^{n+\frac{1}{2}}(i, j, k) - H_z^{n+\frac{1}{2}}(i-1, j, k)}{\delta_x} \right] \quad (5.2b)$$

$$E_z^{n+1}(i, j, k) = P_z(i, j, k)E_z^n(i, j, k) + Q_z(i, j, k) \left[\frac{H_y^{n+\frac{1}{2}}(i, j, k) - E_y^{n+\frac{1}{2}}(i-1, j, k)}{\delta_x} - \frac{H_x^{n+\frac{1}{2}}(i, j, k) - H_x^{n+\frac{1}{2}}(i, j-1, k)}{\delta_y} \right] \quad (5.2c)$$

where, for the x components:

$$P_x(i, j, k) = \frac{1 - \frac{\delta_t \sigma_x(i, j, k)}{2\epsilon_x(i, j, k)}}{1 + \frac{\delta_t \sigma_x(i, j, k)}{2\epsilon_x(i, j, k)}}$$

$$Q_x(i, j, k) = \frac{\frac{\delta_t}{\epsilon_x(i, j, k)}}{1 + \frac{\delta_t \sigma_x(i, j, k)}{2\epsilon_x(i, j, k)}}$$

and similar for the y and z components. If a material behaves as a perfect conductor, having $\sigma \rightarrow \infty$, it is easily seen that $P \rightarrow 1$ while $Q \rightarrow 0$ and so (5.2) do not really update the \vec{E} field values (and consequently also neither the \vec{H} ones). By starting from an initial zero condition, that means the electric and magnetic fields will remain zero inside the spatial region occupied by a perfect conductor, which is the correct behavior.

An important matter is that of the imposition of the correct inner boundary conditions at the transition between two different media. The continuity condition of the tangential components of \vec{E} and \vec{H} to a surface element is directly enforced by the curl equations themselves — from which the continuity conditions are derived analytically — in the limit $\delta_x, \delta_y, \delta_z \rightarrow 0$, in presence of well behaved solutions (no singularities). So the correct solution is assured if the FDTD algorithm converges.

The *computational complexity* of the FDTD bulk algorithm — the number of floating point operations (Flops) it requires — is very favorable compared to other numerical methods like the Method of Moments or the Finite Element Method, because it does not require any linear system solution, being it completely explicit, with each field component updated by means of its first neighbors. Such complexity is $O(N)$, where N is the number of Yee cells used in the model ($N = N_x \times N_y \times N_z$, the Yee cell numbers along the coordinate axes). It grows only linearly with the size of the model, one of the lowest growth rates attainable.

The finite difference equations (5.1), (5.2) represent the algorithm for the electromagnetic field values updating inside the computational volume. When supplemented with an algorithm, described later in Chapter 6, for the computational volume truncation (the boundary conditions in the CPML layer), with a further algorithm for the plane wave inclusion, described later in Chapter 7, and the one for the frequency domain analysis described in the next Section, it becomes a complete algorithm for the numerical solution of the Maxwell's equations, the FDTD method, described diagrammatically in Fig. 5.5.

5.2 Stability

The stability of the bulk algorithm is strictly connected with the convergence of the FDTD method. In fact, even if the proposed (5.1)- (5.2) scheme is consistent (with second order accuracy), its numerical solution will not necessarily converge to the solution of the linear PDEs system of the Maxwell's curl equations. Convergence means that, $\forall t$ fixed in $[0, T]$, if $\delta_t \rightarrow 0$ and simultaneously $n \rightarrow \infty$ in such a way that $n\delta_t = t$ finite, the

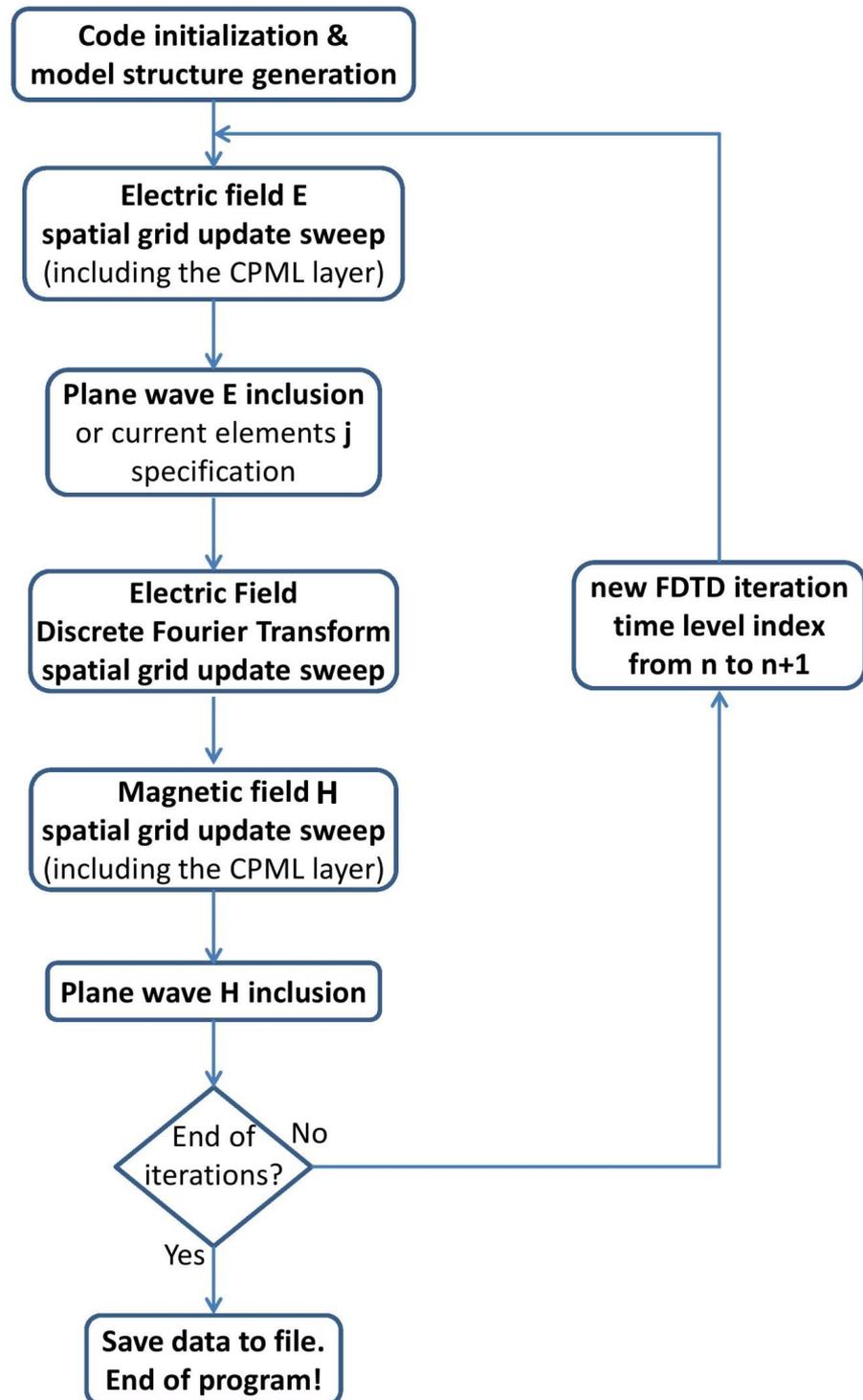


Figure 5.5: Flowchart of the FDTD algorithm.

numerical solution $\vec{E}_{num}, \vec{H}_{num}$ will tend *in norm* to the true solution \vec{E}, \vec{H} if also $\delta_x, \delta_y, \delta_z \rightarrow 0$ with δ_t . The Lax equivalence theorem [20] states that, if the time evolutive IVP is well posed and the finite difference approximation is consistent, stability is the necessary and sufficient condition for convergence. Stability on the other hand, means that the round-off errors due to the necessarily finite accuracy of every computer, have a controlled growth rate in passing from each time level to the next. In this way they will not accumulate — because of the large number of repetitive operations that are involved — during the time marching algorithm. Otherwise, they eventually overwhelm the approximate numerical solution thus making it useless. For the scheme (5.1)- (5.2), the Courant-Friedrichs-Lewy stability condition [13] holds:

$$\delta_t c_{max} \leq \frac{1}{\sqrt{\frac{1}{\delta_x^2} + \frac{1}{\delta_y^2} + \frac{1}{\delta_z^2}}}, \quad (5.3)$$

where c is the phase velocity in a given medium and:

$$c_{max} = \frac{1}{\sqrt{\mu_o \epsilon_{min}}}$$

with ϵ_{min} the smallest electric permittivity (the real part) of the various media involved in the model. Its derivation is found in Appendix C. When, as described in the next Section, frequency domain analyses are performed, the condition (5.3) should be obeyed in the whole spectral frequency interval considered.

5.3 Frequency domain analysis

Despite the FDTD method works in the time-domain, it is possible to perform frequency-domain analyses of the resulting field distributions corresponding to steady harmonic states of a given angular frequency $\omega = 2\pi f$. A direct way to make this would be the use of current density sources varying time-harmonically, or the use of an harmonic plane wave impinging on a target. After a smooth onset (to avoid abrupt changes generating high frequency noise), the excitation signals could effectively start to oscillate sinusoidally with the given frequency. The excitation will propagate inside the physical system and, after a transient interval, all of its parts will oscillate harmonically at the same frequency with a given amplitude \vec{A} and phase ϕ depending on position (limiting amplitude and limiting absorption principles). The amplitude determination would require a routine for holding

the local maximum value at every space location, and the phase determination the local delay — mapped on the $[0, 2\pi]$ interval — measurement with respect to some reference points. That just described is the so called *continuous wave* FDTD method (CW-FDTD). A different more efficient approach in the present work is the *pulsed* FDTD method. It consists in the use of a finite time duration signal as the current density sources or as the profile of a plane wave pulse impinging on a target. The system response will also be concentrated in a finite interval because the excitation is turned off after a while. It is thus possible to perform a frequency analysis of the generic field components at each space location by in-line updating, starting from the first time iteration, its Discrete Fourier Transform (DFT) G , with time step δ_t , approximating the continuous Fourier integral:

$$G(\vec{r}; \omega) \propto \int_{-\infty}^{+\infty} g(\vec{r}, t) e^{i\omega t} dt \Rightarrow G(i, j, k; \omega) \propto \sum_{n=-\infty}^{+\infty} g^n(i, j, k) e^{i\omega n \delta_t}.$$

Actually, only a finite number of terms will contribute to the sum above, because for $n < 0$ or n greater than a certain n_{max} , all components $g^n(i, j, k)$ in the grid are zero. Thus there is no truncation error in the DFT (although there is a less important discretization error due to the time sampling). The angular frequency resolution $\Delta\omega$ of this procedure can be estimated by considering that, to avoid aliasing:

$$\omega n \delta_t \approx 2\pi,$$

from which $n \approx 2\pi/\omega\delta_t$. By differentiating with respect of ω and n the above expression, after solving for $\Delta\omega$, introducing the value for n just obtained, remembering that $\Delta n \approx 1$ and neglecting an unimportant change of sign, one gets:

$$\boxed{\Delta\omega \approx \frac{\omega^2 \delta_t}{2\pi}}.$$

G is a complex phasor bearing information about both the amplitude and the phase of a field component (electric or magnetic) at each spatial grid location i, j, k .

The pulsed FDTD method allows a simultaneous multifrequency analysis with a single run of the FDTD program code. It suffices to have RAM enough to memorize at each spatial location the DFTs for all the frequencies considered. With the CW-FDTD method one would be obliged to make a run for each one of the frequencies analyzed. This is of importance with large

simulations, which take a lot of execution time each.

Examples of finite duration signals (compact pulses) used in the present work are [28, 29]:

$$S(t) = \begin{cases} A_n [1 - \cos(\frac{2\pi t}{T})]^n & \text{if } 0 \leq t \leq T \\ 0 & \text{elsewhere} \end{cases} \quad (5.4)$$

with $n = 1, 2, 3$ and where A_n is a normalization factor. Graphs of the pulses and their Fourier transforms are given in Fig. 5.6 and Fig. 5.7:

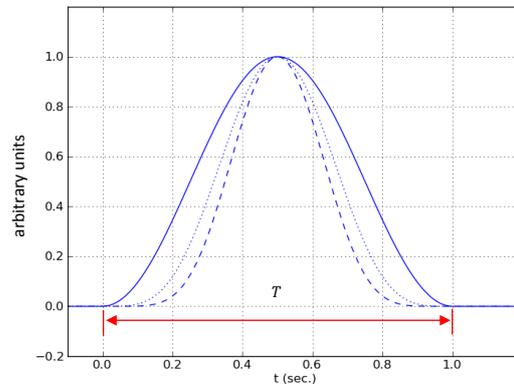


Figure 5.6: Graphs of pulses (5.4) with $n = 1$ (solid line), $n = 2$ (dotted line) and $n = 3$ (dashed line). Normalized heights.

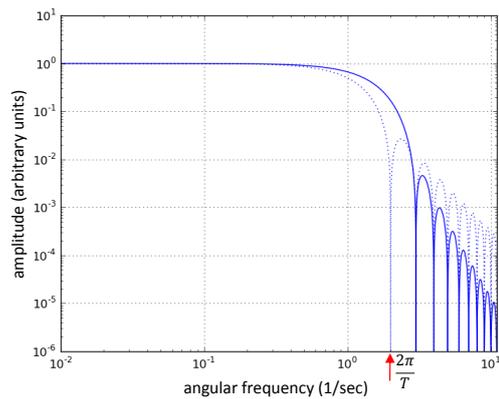


Figure 5.7: Amplitude spectra of the pulses (5.4).

Further examples are given by [28, 29]:

$$S(t) = \begin{cases} A_n \sin\left(\frac{\pi t}{T}\right) \left[1 - \cos\left(\frac{\pi t}{T}\right)\right]^n & \text{if } 0 \leq t \leq 2T \\ 0 & \text{elsewhere} \end{cases} \quad (5.5)$$

with $n = 1, 2$ and, again, where A_n is a normalization factor. Graphs of the pulses and their Fourier transforms are given in Fig. 5.8 and Fig. 5.9:

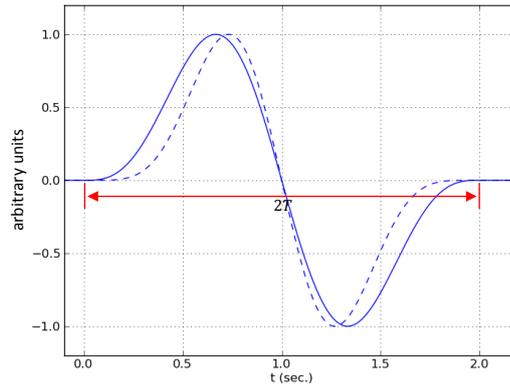


Figure 5.8: Graphs of pulses (5.5) with $n = 1$ (solid line) and $n = 2$ (dashed line). Normalized heights.

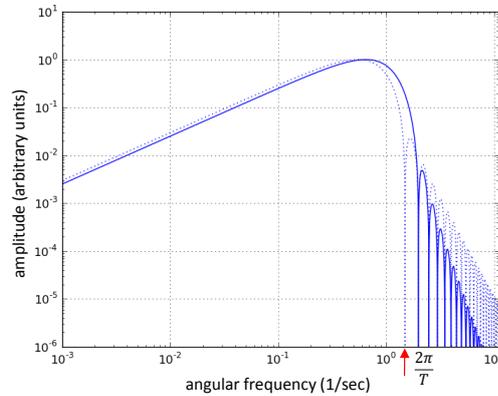


Figure 5.9: Amplitude spectra of the pulses (5.5).

The second example (5.5), a signal with zero mean value, and thus with an amplitude spectrum going to 0 for $\omega \rightarrow 0$ (see Fig. 5.9), could effectively be more suitable if current density sources were present, because it does not produce a permanent electric charge separation at the ends of the current element itself after it has ceased. The first signal pulse (5.4) would indeed

produce an electric dipole inside the spatial computational domain (with a charge separation of δ_x or δ_y or δ_z , depending on the orientation of the current element) whose static field would remain as the trailing part of a transitory onset after the pulse (5.4) has ceased (and contrary to common sense, because a time domain algorithm is eventually not retained to be able to reproduce the static behavior for time independent systems and boundary conditions) [30].

From the above examples and the Indetermination Principle for the Fourier transform: $\Delta\omega\Delta t \approx 2\pi$, it is seen that the pulse duration T has to be fixed well above the maximum frequency one wants to analyze by means of the DFT, because T determines the bandwidth of the exciting signal, beyond which the single frequency amplitude content becomes smaller and smaller, and potentially overwhelmed by numerical noise. Once T has been chosen and δ_t has been fixed according to (5.3), one has the minimum number of FDTD iterations that have to be performed by the code. Effectively the overall number of FDTD iterations has to guarantee enough time for the pulse sollecitation to reach any point inside the spatial grid and to act there for its entire duration. As a rule of thumb, one can evaluate the time employed by the leading edge of the pulse to cover a distance equal to a diagonal of the spatial computational domain, to double it (for a front back to front passage) and then to add T . In evaluating the time to cover a given distance, the minimum (estimated) phase velocity determined by the media inside the target should be used.

Chapter 6

Absorbing boundary conditions

Due to the limited computational resources, only a finite amount of memory (RAM) can be allocated for the \vec{E} and \vec{H} fields values of the spatial computational grid described in Chapter 5. In other words, the bulk algorithm represented by the updating finite-difference equations (5.1) and (5.2) cannot be applied over the entire spatial grid domain, because the components' interlacing, due to their staggering according to the Yee sampling, would require increasingly distant values which are not at all available. Thus, if the computational domain must have a finite spatial extent, it means that at its outer surface suitable *boundary conditions* must be applied to the *tangential field components*, to truncate in some way the bulk algorithm. The action of such boundary conditions has to be the one of isolating the convex modeled system, entirely enclosed in the computational domain, from any outer influence. This substantially means: to prevent any outgoing radiation, generated by the interaction of the sources with the system, from being scattered back in the computational domain, in such a way that the modeled volume can be ideally considered of infinite extent. For this reason in the literature they are known as Absorbing Boundary Conditions (ABC), although in the past they were also called Radiation Boundary Conditions (RBC). Different analytical approaches were used to implement them [21–25]. Nowadays, the most recent trend [34] about the FDTD spatial grid truncation makes use of an *external layer*, acting as a lossy, anisotropic, dispersive material, interfaced at the outer surface of the computational domain. This layer is able to perfectly match the transmission characteristics of the computational volume inner media, thus absorbing any outgoing plane waves without the generation, at the interface, of the corresponding reflected ones. This kind of boundary conditions for the FDTD spatial mesh truncation are described in the next Section. From a mathematical/numerical point of view the Absorbing Boundary Conditions have to guarantee the well-posedness of the PDEs

problem, i.e., do not alter the stability assured by (5.3). In the past this required an in deep analysis [26, 27]. In the present work the well-posedness of the ABCs implemented is assumed on a physical ground.

6.1 Convolutional-Perfectly Matched Layer

To get a physical insight about the way in which a perfectly matched layer (PML) operates, it is useful to consider a one-dimensional propagation source free problem (of the type described in Chapter 4). A plane interface is considered between two media. The one on the left (identified by a subscript 1) is a perfect dielectric with absolute parameters ϵ_1 and μ_1 . The one on the right (identified by a subscript 2) is a lossy electric *and* magnetic material with absolute parameters ϵ_2 , μ_2 and the conductivities σ_2^e and σ_2^m . These conductivities, measured in Ω^{-1}/m and Ω/m respectively, describe losses due to ohmic currents. The magnetic one being formally introduced by duality of the electric case:

$$\vec{\nabla} \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} - \sigma^m \vec{H} \quad (6.1a)$$

$$\vec{\nabla} \times \vec{H} = \epsilon \frac{\partial \vec{E}}{\partial t} + \sigma^e \vec{E}. \quad (6.1b)$$

After passing to the frequency domain by means of the usual temporal imaginary complex exponential $e^{-i\omega t}$, one has that the lossy material to the right is characterizable by means of a complex electric permittivity and a complex magnetic permeability (here denoted by the subscript c for distinguishing between real and complex ones):

$$\begin{aligned} \epsilon_{c2} &= \epsilon_2 + i \frac{\sigma_2^e}{\omega} \\ \mu_{c2} &= \mu_2 + i \frac{\sigma_2^m}{\omega}. \end{aligned}$$

After calculating the reflectance R for the *normal incidence* of a plane wave, by calculations similar to the ones described in Chapter 4, one has:

$$R = \frac{\eta_{c2} - \eta_1}{\eta_{c2} + \eta_1}$$

where:

$$\eta_1 = \sqrt{\frac{\mu_1}{\epsilon_1}}$$

$$\eta_{c_2} = \sqrt{\frac{\mu_2 \left(1 + i \frac{\sigma_2^m}{\omega \mu_2}\right)}{\epsilon_2 \left(1 + i \frac{\sigma_2^e}{\omega \epsilon_2}\right)}}.$$

One thus sees that, by choosing the *perfect matching* conditions:

$$\boxed{\frac{\mu_2}{\mu_1} = \frac{\epsilon_2}{\epsilon_1} \quad \text{and} \quad \frac{\sigma_2^m}{\mu_1} = \frac{\sigma_2^e}{\epsilon_1}} \quad (6.2)$$

there results, for all ω :

$$\eta_{c_2} = \eta_1$$

and, as a consequence, $R = 0$ identically, i.e., no reflection at all for any frequency. Thus, a sufficiently thick lossy electric and magnetic material could act as a perfect absorber because, after the wave has passed through the interface without generating any reflection, it will enter the lossy medium being exponentially damped to a negligible amplitude (and further attenuated if it propagates back toward the lossless region). Unfortunately, such a perfect matching is not attainable when oblique incidence is considered, because an angular dependence appears which prevents the vanishing of the reflectance R at the various outgoing directions. For an arbitrary angle of incidence one has to consider an anisotropic material in a three-dimensional setting [32]. In the frequency domain ($e^{-i\omega t}$) the Maxwell's curl equations become:

$$\vec{\nabla} \times \vec{E} = i\omega \bar{\bar{\mu}} \cdot \vec{H} - \bar{\bar{\sigma}}^m \cdot \vec{H}$$

$$\vec{\nabla} \times \vec{H} = -i\omega \bar{\bar{\epsilon}} \cdot \vec{E} + \bar{\bar{\sigma}}^e \cdot \vec{E}$$

where \vec{E} and \vec{H} depend on the position \vec{r} and the angular frequency ω . Here the double overbar denotes a tensor written as a dyadic (see also Section 2.3 and Appendix A), which justifies the dot product \cdot for the action (from the left) of the tensor on a field vector (which returns a vector). By considering *diagonal* tensors only, and with complex components, the full set of Maxwell's equations (including those with the divergences) can be rewritten as (with a further abuse of notation regarding the tensors):

$$\vec{\nabla} \cdot \bar{\bar{\epsilon}} \cdot \vec{E} = 0$$

$$\vec{\nabla} \cdot \bar{\bar{\mu}} \cdot \vec{H} = 0$$

$$\vec{\nabla} \times \vec{E} = i\omega \bar{\bar{\mu}} \cdot \vec{H}$$

$$\vec{\nabla} \times \vec{H} = -i\omega \bar{\bar{\epsilon}} \cdot \vec{H}$$

where:

$$\bar{\bar{\epsilon}} = \begin{bmatrix} \epsilon_x + i\frac{\sigma_x^e}{\omega} & 0 & 0 \\ 0 & \epsilon_y + i\frac{\sigma_y^e}{\omega} & 0 \\ 0 & 0 & \epsilon_z + i\frac{\sigma_z^e}{\omega} \end{bmatrix}$$

and

$$\bar{\bar{\mu}} = \begin{bmatrix} \mu_x + i\frac{\sigma_x^m}{\omega} & 0 & 0 \\ 0 & \mu_y + i\frac{\sigma_y^m}{\omega} & 0 \\ 0 & 0 & \mu_z + i\frac{\sigma_z^m}{\omega} \end{bmatrix}.$$

By choosing from the outset the *perfect matching* between the anisotropic material and the free space, as suggested by the one-dimensional case above (with the material with index 1 being the free space):

$$\boxed{\begin{matrix} \bar{\bar{\mu}} \\ \mu_o \end{matrix}} = \boxed{\begin{matrix} \bar{\bar{\epsilon}} \\ \epsilon_o \end{matrix}}, \quad (6.3)$$

one can write:

$$\bar{\bar{\epsilon}} = \epsilon_o \bar{\bar{\Lambda}}$$

$$\bar{\bar{\mu}} = \mu_o \bar{\bar{\Lambda}}$$

with

$$\bar{\bar{\Lambda}} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix},$$

where a , b , c are complex numbers. By inserting these expressions in the Maxwell's equations, and trying the usual "ansatz" for plane waves:

$$\vec{E} = \vec{\mathcal{E}} e^{i\vec{k} \cdot \vec{r}}$$

$$\vec{H} = \vec{\mathcal{H}} e^{i\vec{k} \cdot \vec{r}}$$

with $\vec{\mathcal{E}}$, $\vec{\mathcal{H}}$ and \vec{k} constant vectors, one gets:

$$\vec{k} \cdot \bar{\bar{\Lambda}} \cdot \vec{\mathcal{E}} = 0 = \vec{k} \cdot \bar{\bar{\Lambda}} \cdot \vec{\mathcal{H}}$$

$$\vec{k} \times \vec{\mathcal{E}} = \omega \mu_o \bar{\bar{\Lambda}} \cdot \vec{\mathcal{H}}$$

$$\vec{k} \times \vec{\mathcal{H}} = -\omega \epsilon_o \bar{\bar{\Lambda}} \cdot \vec{\mathcal{E}}.$$

By considering the following transformation to new vectors $\vec{\mathcal{E}}'$, $\vec{\mathcal{H}}'$, \vec{k}' :

$$\vec{\mathcal{E}}' = \bar{\bar{\Lambda}}^{1/2} \cdot \vec{\mathcal{E}}$$

$$\vec{\mathcal{H}}' = \bar{\bar{\Lambda}}^{1/2} \cdot \vec{\mathcal{H}}$$

$$\vec{k}' = \frac{1}{\sqrt{abc}} \bar{\bar{\Lambda}}^{1/2} \cdot \vec{k}$$

where:

$$\bar{\bar{\Lambda}}^{1/2} = \begin{bmatrix} \sqrt{a} & 0 & 0 \\ 0 & \sqrt{b} & 0 \\ 0 & 0 & \sqrt{c} \end{bmatrix}$$

(being $\bar{\bar{\Lambda}}$ diagonal), it is possible to discard the occurrence of the tensor (care must be paid with the vector product \times and the $\bar{\bar{\Lambda}}$ tensor: they cannot be simply interchanged):

$$\vec{k}' \cdot \vec{\mathcal{E}}' = 0 = \vec{k}' \cdot \vec{\mathcal{H}}'$$

$$\vec{k}' \times \vec{\mathcal{E}}' = \omega \mu_o \vec{\mathcal{H}}'$$

$$\vec{k}' \times \vec{\mathcal{H}}' = -\omega \epsilon_o \vec{\mathcal{E}}'.$$

It follows that:

$$\vec{k}' \cdot \vec{k}' = \omega^2 \epsilon_o \mu_o = k_o^2$$

like in free space. However, by expressing \vec{k}' as a function of the original \vec{k} one gets:

$$\frac{k_x^2}{bc} + \frac{k_y^2}{ac} + \frac{k_z^2}{ab} = k_o^2.$$

Thus, after introducing a spherical coordinate system with colatitude θ measured from the positive y -semiaxis (which represents the propagation direction) and the azimuthal angle ϕ measured from the positive x -semiaxis, the solution is:

$$k_x = k_o \sqrt{bc} \sin \theta \cos \phi$$

$$k_y = k_o \sqrt{ac} \cos \theta$$

$$k_z = k_o \sqrt{ab} \sin \theta \sin \phi.$$

If a plane wave is now impinging from the free space on an interface with such an anisotropic material, assuming the xy -plane as the plane of incidence ($\phi = 0$) and the xz -plane as the interface between the two media, for a TE wave (\vec{E} along the z -axis normal to the plane of incidence) one has:

$$\vec{E}_{inc} = \hat{z} \mathcal{E} e^{ik_o(x \sin \theta_{inc} + y \cos \theta_{inc})}$$

$$\vec{E}_{refl} = \hat{z} r^{te} \mathcal{E} e^{ik_o(x \sin \theta_{refl} - y \cos \theta_{refl})}$$

$$\vec{E}_{tr} = \hat{z} t^{te} \mathcal{E} e^{ik_o(x \sqrt{bc} \sin \theta_{tr} + y \sqrt{ac} \cos \theta_{tr})}$$

where in the transmitted field, the wave vector is as has been calculated above. The magnetic field can be deduced from the $\vec{\nabla} \times \vec{E}$ Maxwell's equation (including $\bar{\bar{\Lambda}}$ in the anisotropic half-space):

$$\vec{H}_{inc} = \sqrt{\frac{\epsilon_o}{\mu_o}} \mathcal{E} (\hat{x} \cos \theta_{inc} - \hat{y} \sin \theta_{inc}) e^{ik_o(x \sin \theta_{inc} + y \cos \theta_{inc})}$$

$$\vec{H}_{refl} = -\sqrt{\frac{\epsilon_o}{\mu_o}} r^{te} \mathcal{E} (\hat{x} \cos \theta_{refl} + \hat{y} \sin \theta_{refl}) e^{ik_o(x \sin \theta_{refl} - y \cos \theta_{refl})}$$

$$\vec{H}_{tr} = \sqrt{\frac{\epsilon_o}{\mu_o}} t^{te} \mathcal{E} \left(\hat{x} \sqrt{\frac{c}{a}} \cos \theta_{tr} - \hat{y} \sqrt{\frac{c}{b}} \sin \theta_{tr} \right) e^{ik_o(x \sqrt{bc} \sin \theta_{tr} + y \sqrt{ac} \cos \theta_{tr})}.$$

By applying the *phase* matching conditions along the x -axis (the transverse direction) one gets:

$$\sin \theta_{inc} = \sin \theta_{refl} \quad \text{and} \quad \sin \theta_{inc} = \sqrt{bc} \sin \theta_{tr},$$

while for the reflection amplitude r^{te} , by equating the transverse components of \vec{E} and \vec{H} on both sides:

$$r^{te} = \frac{\cos \theta_{inc} - \sqrt{\frac{c}{a}} \cos \theta_{tr}}{\cos \theta_{inc} + \sqrt{\frac{c}{a}} \cos \theta_{tr}}.$$

A similar calculation for the TM case (\vec{H} along the z -axis normal to the plane of incidence) would result simply in a change of sign (but now r^{tm} would be a ratio of magnetic field amplitudes):

$$r^{tm} = \frac{\sqrt{\frac{c}{a}} \cos \theta_{tr} - \cos \theta_{inc}}{\cos \theta_{inc} + \sqrt{\frac{c}{a}} \cos \theta_{tr}}$$

with obviously the same phase matching conditions. Thus, by choosing $bc = 1$, which implies $\theta_{inc} = \theta_{refl} = \theta_{tr}$, and $\frac{c}{a} = 1$, it results: $r^{te} = r^{tm} = 0$, i.e., there are no reflected waves irrespective of the direction of incidence. Summarizing:

$$\boxed{a = \frac{1}{b} = c} \quad (6.4)$$

in $\bar{\Lambda}$ as the further condition, together with (6.3), for getting perfect match between two half-spaces, one of which being the free space [32]. By putting: $a = c = \alpha + i\beta$ ($\beta > 0$), one sees that the imaginary part β contributes to $\sigma_{x,y,z}^e$ and $\sigma_{x,y,z}^m$ and directly determines the exponential decay rate of the transmitted wave along the positive y direction (because $Im\{\sqrt{ac}\} = \beta$, which is the reason to choose it strictly positive).

Another approach in the same vein, but more geometrically oriented, is the one using *stretched coordinates* [33]. It consists in adopting two different forms of a modified $\vec{\nabla}$ operator (in the frequency domain):

$$\vec{\nabla}_e = \hat{x} \frac{1}{e_x} \frac{\partial}{\partial x} + \hat{y} \frac{1}{e_y} \frac{\partial}{\partial y} + \hat{z} \frac{1}{e_z} \frac{\partial}{\partial z}$$

$$\vec{\nabla}_h = \hat{x} \frac{1}{h_x} \frac{\partial}{\partial x} + \hat{y} \frac{1}{h_y} \frac{\partial}{\partial y} + \hat{z} \frac{1}{h_z} \frac{\partial}{\partial z}$$

where e_α and h_α ($\alpha = x, y, z$) are, possibly complex, functions of ω but not of the position for a given homogeneous material. Maxwell's equations are rewritten using the $\vec{\nabla}_e$ and $\vec{\nabla}_h$ operators in a source free medium:

$$\vec{\nabla}_e \times \vec{E} = i\omega \mu \vec{H}$$

$$\vec{\nabla}_h \times \vec{H} = -i\omega \epsilon \vec{E}$$

$$\vec{\nabla}_h \cdot \epsilon \vec{E} = 0$$

$$\vec{\nabla}_e \cdot \mu \vec{H} = 0$$

(the last pair being a consequence of the first pair). With the usual “ansatz” for plane wave solutions in a homogeneous medium: $\vec{E} = \vec{\mathcal{E}} e^{i\vec{k} \cdot \vec{r}}$ and $\vec{H} = \vec{\mathcal{H}} e^{i\vec{k} \cdot \vec{r}}$, with $\vec{k} = \hat{x}k_x + \hat{y}k_y + \hat{z}k_z$ ($\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$ being constant vectors), from the curl equations it follows that:

$$\vec{k}_e \times \vec{\mathcal{E}} = \omega \mu \vec{\mathcal{H}} \quad (6.5)$$

$$\vec{k}_h \times \vec{\mathcal{H}} = -\omega \epsilon \vec{\mathcal{E}} \quad (6.6)$$

where:

$$\vec{k}_e = \hat{x} \frac{k_x}{e_x} + \hat{y} \frac{k_y}{e_y} + \hat{z} \frac{k_z}{e_z} \quad (6.7)$$

$$\vec{k}_h = \hat{x} \frac{k_x}{h_x} + \hat{y} \frac{k_y}{h_y} + \hat{z} \frac{k_z}{h_z}. \quad (6.8)$$

Left multiplying (6.6) by \vec{k}_e :

$$\vec{k}_e \times \vec{k}_h \times \vec{\mathcal{H}} = -\omega^2 \mu \epsilon \vec{\mathcal{H}} = \vec{k}_h (\vec{k}_e \cdot \vec{\mathcal{H}}) - \vec{\mathcal{H}} (\vec{k}_e \cdot \vec{k}_h),$$

where use has made of the *BAC*–*CAB* identity for the triple vector product. Taking account that, from (6.5), $\vec{k}_e \cdot \vec{\mathcal{H}} = 0$, one gets the constraint:

$$\omega^2 \mu \epsilon = \vec{k}_e \cdot \vec{k}_h = \frac{1}{e_x h_x} k_x^2 + \frac{1}{e_y h_y} k_y^2 + \frac{1}{e_z h_z} k_z^2,$$

which, after considering the same spherical coordinate system previously introduced, is satisfied if:

$$k_x = \omega \sqrt{\mu \epsilon} \sqrt{e_x h_x} \sin \theta \cos \phi$$

$$k_y = \omega \sqrt{\mu \epsilon} \sqrt{e_y h_y} \cos \theta$$

$$k_z = \omega \sqrt{\mu \epsilon} \sqrt{e_z h_z} \sin \theta \sin \phi.$$

Then with an interface, represented by the plane $y = 0$, between two media identified by the subscripts 1, for the $y < 0$ half-space and 2, one can write

for the TE case ($\vec{E} = \mathcal{E}\hat{u}$ with \hat{u} a unit vector in the transverse direction, i.e. contained in the xz plane) of a plane wave propagating along the positive y axis:

$$\vec{E}_1 = \mathcal{E}\hat{u}e^{i\vec{k}_{inc}\cdot\vec{r}} + r^{te}\mathcal{E}\hat{u}e^{i\vec{k}_{refl}\cdot\vec{r}}$$

$$\vec{E}_2 = t^{te}\mathcal{E}\hat{u}e^{i\vec{k}_{tr}\cdot\vec{r}}$$

for the electric field on either side and, from (6.5):

$$\vec{H}_1 = \mathcal{E}\frac{\vec{k}_{e,inc} \times \hat{u}}{\omega\mu_1}e^{i\vec{k}_{inc}\cdot\vec{r}} + r^{te}\mathcal{E}\frac{\vec{k}_{e,refl} \times \hat{u}}{\omega\mu_1}e^{i\vec{k}_{refl}\cdot\vec{r}}$$

$$\vec{H}_2 = t^{te}\mathcal{E}\frac{\vec{k}_{e,tr} \times \hat{u}}{\omega\mu_2}e^{i\vec{k}_{tr}\cdot\vec{r}}$$

for the magnetic one. Here r^{te} and t^{te} are the reflection and transmission electric field amplitudes in the TE case. In the above expressions for the fields, $\vec{k}_{e,inc}$, $\vec{k}_{e,refl}$, $\vec{k}_{e,tr}$ are the corresponding \vec{k}_{inc} , \vec{k}_{refl} , \vec{k}_{tr} in the complex exponentials (in the phases) but “corrected” according to (6.7). Obviously all these wave vectors have an y component along the impinging direction. Moreover, it is:

$$(\vec{k}_{refl})_y = -(\vec{k}_{inc})_y$$

and

$$(\vec{k}_{refl})_x = (\vec{k}_{inc})_x \quad (\vec{k}_{refl})_z = (\vec{k}_{inc})_z$$

for *phase matching* and, consequently, similar relations for $\vec{k}_{e,inc}$ and $\vec{k}_{e,refl}$. By equating the *tangential* components of \vec{E} and \vec{H} on either side of the interface (a continuity condition following from the “modified” $\vec{\nabla}_e$ and $\vec{\nabla}_h$ curl equations in the usual manner) and using now media subscripts 1 and 2 instead of the *inc*, *refl* and *tr* subscripts, one gets:

$$\omega\sqrt{\epsilon_1\mu_1}\sqrt{e_{x,1}h_{x,1}}\sin\theta_1\cos\phi_1 = \omega\sqrt{\epsilon_2\mu_2}\sqrt{e_{x,2}h_{x,2}}\sin\theta_2\cos\phi_2$$

$$\omega\sqrt{\epsilon_1\mu_1}\sqrt{e_{z,1}h_{z,1}}\sin\theta_1\sin\phi_1 = \omega\sqrt{\epsilon_2\mu_2}\sqrt{e_{z,2}h_{z,2}}\sin\theta_2\sin\phi_2$$

(they correspond to $k_{x,1} = k_{x,2}$ and $k_{z,1} = k_{z,2}$ in the field phases) for further phase matching, which allows one to cancel out from each term the exponential factors, and:

$$1 + r^{te} = t^{te}$$

$$\sqrt{\frac{\epsilon_1}{\mu_1}} \sqrt{\frac{h_{y,1}}{e_{y,1}}} \cos \theta_1 (1 - r^{te}) = t^{te} \sqrt{\frac{\epsilon_2}{\mu_2}} \sqrt{\frac{h_{y,2}}{e_{y,2}}} \cos \theta_2$$

which, when solved for r^{te} give:

$$r^{te} = \frac{\sqrt{\frac{\epsilon_1}{\mu_1}} \sqrt{\frac{h_{y,1}}{e_{y,1}}} \cos \theta_1 - \sqrt{\frac{\epsilon_2}{\mu_2}} \sqrt{\frac{h_{y,2}}{e_{y,2}}} \cos \theta_2}{\sqrt{\frac{\epsilon_1}{\mu_1}} \sqrt{\frac{h_{y,1}}{e_{y,1}}} \cos \theta_1 + \sqrt{\frac{\epsilon_2}{\mu_2}} \sqrt{\frac{h_{y,2}}{e_{y,2}}} \cos \theta_2}.$$

If one considers the TM case ($\vec{H} = \mathcal{H}\hat{u}$ with \hat{u} a unit vector in the transverse direction, i.e. contained in the xz plane) and uses (6.6) and (6.8), for the magnetic field reflection amplitude r^{tm} would get:

$$r^{tm} = \frac{\sqrt{\frac{\mu_1}{\epsilon_1}} \sqrt{\frac{e_{y,1}}{h_{y,1}}} \cos \theta_1 - \sqrt{\frac{\mu_2}{\epsilon_2}} \sqrt{\frac{e_{y,2}}{h_{y,2}}} \cos \theta_2}{\sqrt{\frac{\mu_1}{\epsilon_1}} \sqrt{\frac{e_{y,1}}{h_{y,1}}} \cos \theta_1 + \sqrt{\frac{\mu_2}{\epsilon_2}} \sqrt{\frac{e_{y,2}}{h_{y,2}}} \cos \theta_2}.$$

Now, if there is a perfect match expressed by the choices:

$$\boxed{e_x = h_x \quad e_y = h_y \quad e_z = h_z,}$$

$$\frac{\mu_1}{\mu_2} = \frac{\epsilon_1}{\epsilon_2}$$

(already seen) and:

$$\boxed{e_{x,1} = e_{x,2} \quad e_{z,1} = e_{z,2},}$$

one has $\theta_1 = \theta_2$ and $\phi_1 = \phi_2$ from the phase matching conditions but then, also, $r^{te} = r^{tm} = 0$ identically, whatever is the angular frequency ω and *irrespectively from the angle of incidence*. The remaining degree of freedom is the choice:

$$e_{y,2} = h_{y,2} = s_y(\omega) \tag{6.9}$$

with $s_y(\omega)$ a complex valued function of ω , in such a way to have attenuation in the y direction for a wave coming from the non-absorptive medium 1 ($e_{y,1} = h_{y,1} = 1$) and impinging on medium 2.

The preceding analysis allows the theoretical realization of a perfectly

matched medium, able to absorb plane waves originating from a conventional material and propagating along a given coordinate direction, without generating reflection at the interface. One can then think to layers of such a medium terminating an FDTD computational volume at his six faces, thus forming well suited and effective material Absorbing Boundary Conditions by means of which truncate the FDTD spatial grid. They are examples of a Perfectly Matched Layer (PML), as it is referred to in literature. The step further consists in traslating the above formulation from the frequency to the time domain, in such a way it can be useful for FDTD numerical purposes. To this end the time domain Maxwell's equations (1.1c) and (1.1d) are now reconsidered in the light of the stretched coordinate formalism for a PML medium, assuming that in a three dimensional FDTD model, perfect matching and absorption are required for all the three cartesian directions x , y and z . It is considered a single form of the stretched $\vec{\nabla}_s$ operator which, according with the choice (6.9), but in view of the fact it must absorb along all the three cartesian directions, will be written as:

$$\vec{\nabla}_s = \hat{x} \frac{1}{s_x} \frac{\partial}{\partial x} + \hat{y} \frac{1}{s_y} \frac{\partial}{\partial y} + \hat{z} \frac{1}{s_z} \frac{\partial}{\partial z}.$$

where [34, 35]

$$s_r(\omega) = \kappa_r + \frac{\sigma_r}{\alpha_r - i\omega\epsilon_o} \quad (6.10)$$

for $r = x, y, z$. Here κ_r , α_r , σ_r are real parameters (α_r and σ_r strictly positive and having units of $\Omega^{-1}m^{-1}$, $\kappa_r \geq 1$ and adimensional). Their significance is purely geometrical in the context of the stretched coordinate formulation, but obviously the functional form of s_r with respect to the angular frequency ω is physically suggested by the fact it enters in the wavenumber expression for the complex phase of the wave, and has thus to determine an amplitude attenuation like any other conventional material having a dissipation σ . The novelty is that by means of the streched coordinate formulation it is now possible to avoid any reflections. The form of $s_r(\omega)$ is also suggested by comparing it with the entries in the expression for $\bar{\epsilon}$ previously given for the tensorial formulation of a PML medium. Also, the fact that $s_r(\omega)$ is strictly connected with the constitutive parameters of a material requires, by causality, that its real and imaginary parts be coupled by Kramers-Kronig dispersion relations, as is emphasized in [36]. Moreover, complex conjugation of $s_r(\omega)$ has to amount to a change of sign of ω : $s_r^*(\omega) = s_r(-\omega)$. These requests contribute to limit the arbitrariness in the choice of the functional form of $s_r(\omega)$. The main variation in (6.10) is the introduction of the term α_r in the denominator: it is an off the origin pole shifting [36] into

the lower-half complex plane to avoid the singularity at low frequencies, a circumstance which could hurt the low (ideally null) reflectance of a PML layer, as reported in literature [34]. By keeping in mind that the product of $s_r^{-1}(\omega)$ with the spatial derivative of a field component depending on \vec{r} and ω , as results when the operator $\vec{\nabla}_s \times$ is applied to a vector field in Maxwell's curl equations translates, in the time domain, in the *temporal convolution* between the *inverse Fourier transform* \mathcal{F}^{-1} of $s_r^{-1}(\omega)$, here denoted as $\tilde{s}_r(t)$, with the same spatial derivative of the field component as a function of \vec{r} and t (i.e., of the inverse Fourier transform of that field component). To evaluate $\tilde{s}_r(t)$ one should observe that:

$$\lim_{\omega \rightarrow \infty} s_r^{-1}(\omega) = \frac{1}{\kappa_r},$$

thus $\tilde{s}_r(t)$ cannot be evaluated by means of a conventional Fourier integral. However, in distributional sense:

$$\mathcal{F}^{-1}\left(\frac{1}{\kappa_r}\right)(t) \propto \frac{\delta(t)}{\kappa_r}$$

where $\delta(t)$ is the Dirac delta function. By removing the singularity one can then write:

$$\mathcal{F}^{-1}\left(s_r^{-1}(\omega) - \frac{1}{\kappa_r}\right) = \frac{\sigma_r}{i\epsilon_o\kappa_r^2} \int_{-\infty}^{+\infty} \frac{e^{-i\omega t}}{\omega + \frac{i}{\epsilon_o}\left(\frac{\sigma_r}{\kappa_r} + \alpha_r\right)} d\omega,$$

in which a choice has been made for the normalization factor in the direct/inverse Fourier transform pair (it is set to unity). The integral can be evaluated by means of a complex plane contour integral, on a path consisting of a straight line and a semicircle. The integrand has a simple pole lying in the lower half plane (with the choice made above for σ_r , κ_r and α_r). If $t > 0$ the path can be closed from below (the semicircle is in the lower complex half plane) encircling the pole and traversed clockwise. If $t < 0$ the path can be closed from above and oriented counterclockwise, but the integrand, being analytic in that region, gives no contribution at all. Thus:

$$\tilde{s}_r(t) = \frac{2\pi\delta(t)}{\kappa_r} - \frac{2\pi\sigma_r}{\epsilon_o\kappa_r^2} \theta(t) e^{-\frac{1}{\epsilon_o}\left(\frac{\sigma_r}{\kappa_r} + \alpha_r\right)t} = \frac{2\pi\delta(t)}{\kappa_r} + \zeta_r(t) \quad (6.11)$$

where $\theta(t)$ is the unit step distribution and $r = x, y, z$. The Maxwell's equations (6.1) for the PML medium are now written for the x components as:

$$-\mu \frac{\partial H_x}{\partial t} - \sigma^m = \tilde{s}_y * \frac{\partial E_z}{\partial y} - \tilde{s}_z * \frac{\partial E_y}{\partial z}$$

$$\epsilon \frac{\partial E_x}{\partial t} + \sigma^e = \tilde{s}_y * \frac{\partial H_z}{\partial y} - \tilde{s}_z * \frac{\partial H_y}{\partial z}$$

and similar for the y and z components. Here $a * b$ means temporal convolution of two functions of time t :

$$(a * b)(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} a(t - \tau)b(\tau)d\tau = \frac{1}{2\pi} \int_{-\infty}^{+\infty} a(\tau)b(t - \tau)d\tau$$

where the appearance of the $\frac{1}{2\pi}$ factor is consequent to the choice for the normalization factor in the direct/inverse transform pair made above, in such a way that $\mathcal{F}(a * b)(\omega) = \mathcal{F}(a)\mathcal{F}(b)$ with unity as the normalization factor in each direct Fourier transform (from t to ω). By using (6.11) one gets:

$$-\mu \frac{\partial H_x}{\partial t} - \sigma^m = \left(\frac{1}{\kappa_y} \frac{\partial E_z}{\partial y} - \frac{1}{\kappa_z} \frac{\partial E_y}{\partial z} \right) + \left(\zeta_y * \frac{\partial E_z}{\partial y} - \zeta_z * \frac{\partial E_y}{\partial z} \right) \quad (6.12a)$$

$$\epsilon \frac{\partial E_x}{\partial t} + \sigma^e = \left(\frac{1}{\kappa_y} \frac{\partial H_z}{\partial y} - \frac{1}{\kappa_z} \frac{\partial H_y}{\partial z} \right) + \left(\zeta_y * \frac{\partial H_z}{\partial y} - \zeta_z * \frac{\partial H_y}{\partial z} \right) \quad (6.12b)$$

and similar for the y and z components. To be able to write explicit update finite differences expressions for \vec{E} and \vec{H} , the convolutions are evaluated discretely following a recursive-convolution approach already used in [37], [38] for the FDTD description of temporal dispersive media. The discretized version of the convolution of ζ_r with the generic spatial derivative $\frac{\partial}{\partial r}$ of a generic field component F_s , which could be of \vec{E} or of \vec{H} (a subscript e or h will be also attached later to ψ_r to distinguish them), is written as (indicating only the indexed time dependence and not the spatial ones):

$$\begin{aligned} \psi_r(n) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \zeta_r(\tau) \frac{\partial F_s}{\partial r}(n - \tau)d\tau = \frac{1}{2\pi} \int_0^{n\delta t} \zeta_r(\tau) \frac{\partial F_s}{\partial r}(n - \tau)d\tau = \dots \\ &\dots = \frac{1}{2\pi} \sum_{k=0}^{n-1} \int_{k\delta t}^{(k+1)\delta t} \zeta_r(\tau) \frac{\partial F_s}{\partial r}(n - \tau)d\tau \end{aligned}$$

(no sum implied on α) where the second equality follows from causality — the step function in (6.11) — and the fact that before $t = 0$ all field components are zero, while the third equality in the second line follows from the

additive property of the integral. Here δt means the time step in the FDTD discretization. According to the discrete nature of the fields sampling, the spatial derivative $\frac{\partial}{\partial r}$ under the integral is assumed constant on any subinterval of span δt , which means:

$$\psi_r(n) = \frac{1}{2\pi} \sum_{k=0}^{n-1} \int_{k\delta t}^{(k+1)\delta t} \zeta_r(\tau) \frac{\partial F_s}{\partial r}(n - \tau) d\tau = \sum_{k=0}^{n-1} Z_r(k) \frac{\partial F_s}{\partial r}(n - k)$$

where

$$Z_r(k) = \frac{1}{2\pi} \int_{k\delta t}^{(k+1)\delta t} \zeta_r(\tau) d\tau$$

and where, in view of (6.11), the last integral can be calculated explicitly in closed form:

$$Z_r(k) = \frac{\sigma_r \left[e^{-\frac{1}{\epsilon_o} \left(\frac{\sigma_r}{\kappa_r} + \alpha_r \right) \delta t} - 1 \right]}{\kappa_r (\alpha_r \kappa_r + \sigma_r)} e^{-\frac{1}{\epsilon_o} \left(\frac{\sigma_r}{\kappa_r} + \alpha_r \right) k \delta t} = \mathcal{Z}_r e^{-\frac{1}{\epsilon_o} \left(\frac{\sigma_r}{\kappa_r} + \alpha_r \right) k \delta t}$$

from which it follows immediately that:

$$Z_r(k+1) = Z_r(k) e^{-\frac{1}{\epsilon_o} \left(\frac{\sigma_r}{\kappa_r} + \alpha_r \right) \delta t} = Z_r(k) b_r. \quad (6.13)$$

Now, to calculate the convolution $\psi_r(n)$ it is no necessary to save in memory the past evolution history of the (spatial derivatives of the) field components, because it can be updated dynamically by a simple recursion relation, due to the fact that:

$$\psi_r(n) = \sum_{k=0}^{n-1} Z_r(k) \frac{\partial F_s}{\partial r}(n - k) = \sum_{k=1}^{n-1} Z_r(k) \frac{\partial F_s}{\partial r}(n - k) + \mathcal{Z}_r \frac{\partial F_s}{\partial r}(n),$$

from which it follows, by the index change $k = \ell + 1$, that:

$$\psi_r(n) = \sum_{\ell=0}^{n-2} Z_r(\ell+1) \frac{\partial F_s}{\partial r}(n - 1 - \ell) + \mathcal{Z}_r \frac{\partial F_s}{\partial r}(n)$$

but, by means of (6.13), the sum can be put in such a way that:

$$\psi_r(n) = b_r \psi_r(n-1) + \mathcal{Z}_r \frac{\partial F_s}{\partial r}(n). \quad (6.14)$$

In view of (6.14), to modelize a PML medium one needs to allocate extra memory for the ψ_r quantities, which are then simply updatable, at each space location, using the pre-existing values of ψ_r themselves and those of the electromagnetic field. The FDTD discretization of (6.12) is thus as follows:

$$\begin{aligned}
H_x^{n+\frac{1}{2}}(i, j, k) = & P_{h,x}(i, j, k) H_x^{n-\frac{1}{2}}(i, j, k) + \\
& Q_{h,x}(i, j, k) \left[\frac{E_y^n(i, j, k+1) - E_y^n(i, j, k)}{\kappa_z(i, j, k) \delta_z} - \right. \\
& \left. \frac{E_z^n(i, j+1, k) - E_z^n(i, j, k)}{\kappa_y(i, j, k) \delta_y} + \right. \\
& \left. \psi_{h,z}^n(i, j, k) - \psi_{h,y}^n(i, j, k) \right]
\end{aligned} \tag{6.15a}$$

$$\psi_{h,z}^n = b_z(i, j, k) \psi_{h,z}^{n-1} + \mathcal{Z}_z(i, j, k) \left[\frac{E_y^n(i, j, k+1) - E_y^n(i, j, k)}{\delta_z} \right] \tag{6.15b}$$

$$\psi_{h,y}^n = b_y(i, j, k) \psi_{h,y}^{n-1} + \mathcal{Z}_y(i, j, k) \left[\frac{E_z^n(i, j+1, k) - E_z^n(i, j, k)}{\delta_y} \right] \tag{6.15c}$$

with similar expressions for the $H_y^{n+\frac{1}{2}}$ and $H_z^{n+\frac{1}{2}}$ components, and:

$$\begin{aligned}
E_x^{n+1}(i, j, k) = & P_{e,x}(i, j, k) E_x^n(i, j, k) + \\
& Q_{e,x}(i, j, k) \left[\frac{H_z^{n+\frac{1}{2}}(i, j, k) - H_z^{n+\frac{1}{2}}(i, j-1, k)}{\kappa_y(i, j, k) \delta_y} - \right. \\
& \left. \frac{H_y^{n+\frac{1}{2}}(i, j, k) - H_y^{n+\frac{1}{2}}(i, j, k-1)}{\kappa_z(i, j, k) \delta_z} + \right. \\
& \left. \psi_{e,y}^{n+\frac{1}{2}}(i, j, k) - \psi_{e,z}^{n+\frac{1}{2}}(i, j, k) \right]
\end{aligned} \tag{6.16a}$$

$$\begin{aligned} \psi_{e,y}^{n+\frac{1}{2}} &= b_y(i, j, k) \psi_{e,y}^{n-\frac{1}{2}} + \\ & \mathcal{Z}_y(i, j, k) \left[\frac{H_z^{n+\frac{1}{2}}(i, j, k) - H_z^{n+\frac{1}{2}}(i, j-1, k)}{\delta_y} \right] \end{aligned} \quad (6.16b)$$

$$\begin{aligned} \psi_{e,z}^{n+\frac{1}{2}} &= b_z(i, j, k) \psi_{e,z}^{n-\frac{1}{2}} + \\ & \mathcal{Z}_z(i, j, k) \left[\frac{H_y^{n+\frac{1}{2}}(i, j, k) - H_y^{n+\frac{1}{2}}(i, j, k-1)}{\delta_z} \right] \end{aligned} \quad (6.16c)$$

with similar expressions for the E_y^{n+1} and E_z^{n+1} components. It should be remembered that in (6.15) and (6.16) the indices i , j and k address the Yee cell and that inside the cell the x , y and z components of \vec{E} and \vec{H} are distributed with some half cell's edge shifts according to Figs. 5.1 and 5.2. In the above finite-difference expressions also the coefficients b_r , \mathcal{Z}_r and κ_r bring a spatial dependence because inside the PML medium they vary with the position in a manner specified below. Moreover for the P and Q coefficients:

$$\begin{aligned} P_{h,x}(i, j, k) &= \frac{1 - \frac{\delta_t \sigma_x^h(i, j, k)}{2\mu_o}}{1 + \frac{\delta_t \sigma_x^h(i, j, k)}{2\mu_o}} \\ Q_{h,x}(i, j, k) &= \frac{\frac{\delta_t}{\mu_o}}{1 + \frac{\delta_t \sigma_x^h(i, j, k)}{2\mu_o}} \\ P_{e,x}(i, j, k) &= \frac{1 - \frac{\delta_t \sigma_x^e(i, j, k)}{2\epsilon'}}{1 + \frac{\delta_t \sigma_x^e(i, j, k)}{2\epsilon'}} \\ Q_{e,x}(i, j, k) &= \frac{\frac{\delta_t}{\epsilon'}}{1 + \frac{\delta_t \sigma_x^e(i, j, k)}{2\epsilon'}}. \end{aligned}$$

In the above expressions the magnetic permeability of the PML medium is taken as constant and equal to the vacuum value (non-magnetic materials). The electric permittivity is also taken as the constant ϵ' , but accordingly to the matching condition (6.2) it should be equal to the value of the electric permittivity of the nearby material inside the computational volume which it

terminates. Any inhomogeneities in the outer shell of the embedding medium in the computational volume could be accounted for by varying ϵ' in the PML medium accordingly. Obviously σ_r^m , which is introduced for symmetry considerations after the introduction of the extra absorption due to σ_r^e , must satisfy the perfect matching condition (6.2): $\sigma_r^e \mu_o / \epsilon'$. They typically are both chosen equal to 0, because the absorption is handled by the σ_r parameter in (6.11). In fact, the parameters σ_r , κ_r and α_r ($r = x, y, z$) of (6.11) vary depending on the position along a direction perpendicular to the layer, on a segment of length equal to the layer thickness. To make clear this situation, one has to keep in mind that, if the computational volume V is made of N_x , N_y and N_z Yee cells along the x , y and z coordinate axes, then its edges are $L_x = N_x \delta_x$, $L_y = N_y \delta_y$ and $L_z = N_z \delta_z$ respectively, where δ_x , δ_y and δ_z are the space-step sizes (the edges of the Yee cells). One then considers a volume V' of sides $L'_x = \Delta_x + L_x + \Delta_x$, $L'_y = \Delta_y + L_y + \Delta_y$ and $L'_z = \Delta_z + L_z + \Delta_z$ containing V , centered in it. Viewed V and V' as point sets, one can think at the set complement $V' \setminus V$, which is precisely the PML boundary layer: a shell enclosing the computational volume. Its thickness measured in Yee cells is Δ_x / δ_x along x , Δ_y / δ_y along y and Δ_z / δ_z along z . A two-dimensional schematic is shown in Fig. 6.1:

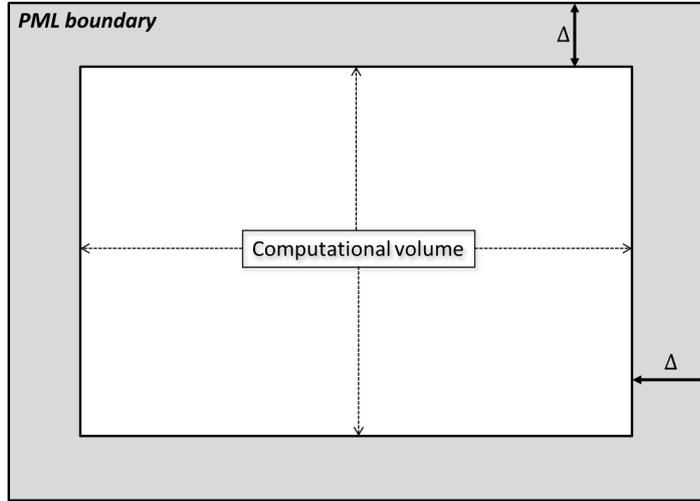


Figure 6.1: Two-dimensional schematic view of the PML boundary layer, with thickness Δ , surrounding the FDTD computational volume.

The PML boundary layer is formed of 6 terminating slabs, of thicknesses $\Delta_{x,y,z}$, which are partially overlapped in correspondence of the parallelepiped

edges and vertices as is suggested by Fig. 6.2:

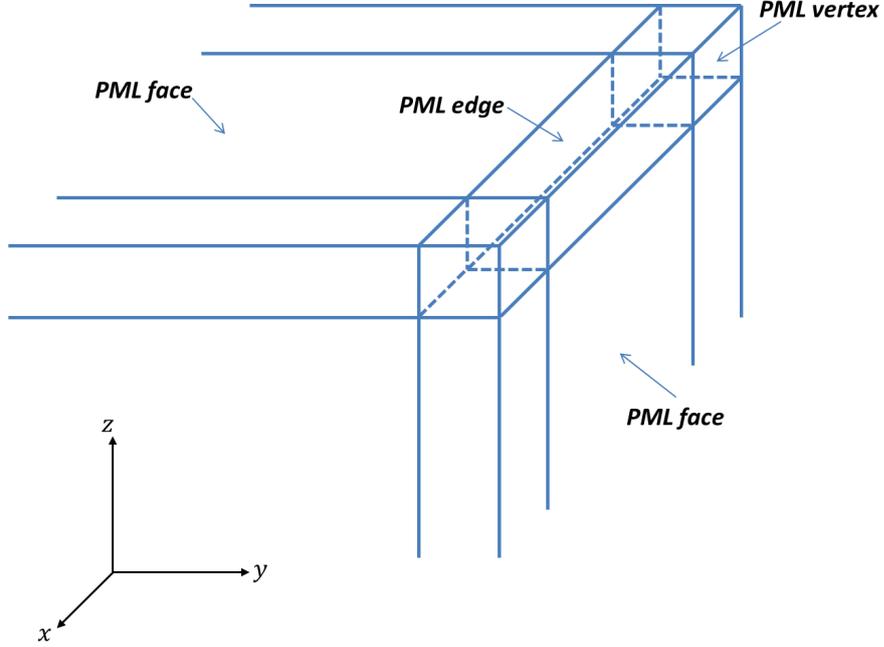


Figure 6.2: Overlap of terminating PML boundary slabs around the FDTD computational volume.

The slabs can be paired along each coordinate direction $r = x, y, z$, and the above parameters vary inside them with respect to a generic coordinate ξ , along that direction, in the interval $[0, \Delta_r]$. According to [39], the laws of variation are empirically chosen to be:

$$\alpha_r(\xi) = \alpha_{r,max} \left(\frac{\Delta_r - \xi}{\Delta_r} \right)^{t_\alpha}$$

$$\sigma_r(\xi) = \sigma_{r,max} \left(\frac{\xi}{\Delta_r} \right)^t$$

$$\kappa_r(\xi) = 1 + \left(\frac{\xi}{\Delta_r} \right)^t (\kappa_{r,max} - 1)$$

($r = x, y, z$) which hold in the increasing coordinate direction: for the coupled opposite slab these behaviors are mirrored about $\xi = 0$. It can be seen that σ_r and κ_r are chosen in such a way to have a smooth onset, but α_r in (6.11) has a decreasing value toward 0. Typically [39]:

$$\kappa_{r,max} = 5$$

$$\alpha_{r,max} = 0.05$$

$$t = 3$$

$$t_\alpha = 1$$

while

$$\sigma_{r,max} = \frac{0.8(t+1)}{Z_o \delta_r \sqrt{\mu_{rel} \epsilon_{rel}}}$$

with Z_o the vacuum impedance ($\approx 377 \Omega$). It is reported in [34, 36, 39] that the α_r parameter in (6.11) governs the absorption of the evanescent fields. It results, from the variability direction of these parameters, which is the one corresponding to the pairing direction of the PML slabs, that the convolutional formalism above described applies only on the electromagnetic field components in the PML medium which are orthogonal to that direction. In fact they exhibit a coefficient variability, in their finite-difference expressions, which corresponds to that direction.

The algorithm described in this section is called *Convolutional* PML (CPML), and has very good performances in the FDTD computational volume truncation.

6.2 Periodic boundary conditions

Periodic boundary conditions (PBC) are not, properly said, absorbing boundary conditions. However, they can be used in conjunction with absorbing boundary conditions, like the CPML ones just described, to simulate an infinite periodic extension along a given coordinate direction. If this is the case, for example, of the ξ axis ($\xi = x, y, z$), then along that direction the CPML algorithm is excluded and the tangential electric field components at the two outer surfaces perpendicular to it, at $\ell = 0$ and $\ell = N_\xi$ (ℓ an index on ξ with N_ξ Yee cells), are calculated simply by equating them to the immediate interior ones of the other side, according to the scheme in Fig. 6.3:

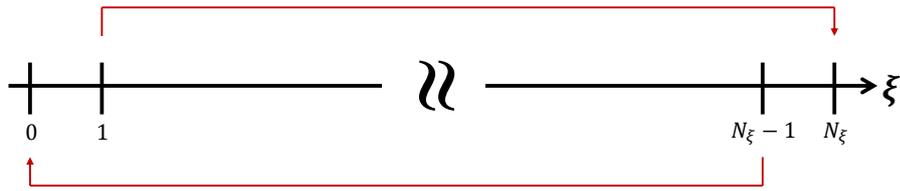


Figure 6.3: Scheme for the tangential electric field completion at periodic boundaries.

For systems whose structure is a periodic extension along one or two perpendicular directions of a given unit, one or two dimensional, “cell”, PBCs allows a substantial saving of computational resources. Obviously in the periodic symmetry of the system has to be included also the incident field distribution. If this is not periodic along the direction/s considered, the PBC cannot be applied.

Chapter 7

Linearly polarized plane wave excitation

In simulations for optical applications, one has to model almost invariably a given material system (the target) and a beam of light interacting with it (primary field). In a time-domain approach as the one used in the present work, *causality* requires to consider the process in which a pulse or a wave train traveling in vacuum (air), with a wavefront moving toward the target and a trailing perturbation, impinges on it. The target's optical properties are then deduced by analyzing the superposition of the primary field with the secondary one, that is the one produced by interaction of the impinging beam with the target itself. It is thus of importance to be able to “inject” a traveling pulse or traveling wave train on the numerical Yee grid previously described in Chapter 5. The simpler space-time dependence of a traveling, linearly polarized, electromagnetic field propagating in vacuum in the direction defined by the unit vector \hat{u} , is that of a plane pulse or of a harmonic plane wave:

$$\vec{E}(\vec{r}, t) = \hat{v} f \left(t - \frac{[\vec{r} - \vec{r}'] \cdot \hat{u}}{c_o} \right) \quad (7.1a)$$

$$\vec{H}(\vec{r}, t) = \frac{1}{Z_o} (\hat{u} \times \hat{v}) f \left(t - \frac{[\vec{r} - \vec{r}'] \cdot \hat{u}}{c_o} \right) \quad (7.1b)$$

where \hat{v} is a unit vector perpendicular to \hat{u} : $\hat{u} \cdot \hat{v} = 0$ (the polarization direction), c_o is the light speed in vacuum, $Z_o = \sqrt{\mu_o/\epsilon_o}$ is the vacuum characteristic impedance, $(\vec{r} - \vec{r}') \cdot \hat{u} = 0$ is the equation defining a plane in \mathbb{R}^3 passing through \vec{r}' and normal to \vec{u} . One could say that it is the starting plane of the wave: all other constant phase plane are parallel to it and on the side indicated by \vec{u} . To this end f has to be a sufficiently well behaved

causal function: $f(t) = 0$ if $t < 0$. One has a plane pulse if f is like S in (5.4) or (5.5) of Chapter 5. If f is a sinusoid restricted to the positive time axis one has a harmonic plane wave. In any case (7.1) are solutions of the Maxwell's equations in a source free region of space and can be used as the primary electromagnetic field, i.e. the light impinging beam. Although there is no need to include the sources producing the plane waves, the real effectiveness of (7.1) stands in the possibility they offer to specify the primary field on a surface completely surrounding the target, without any impractical specification of the field as a three-dimensional distribution at $t = 0$ (as in a pure Initial Value Problem). Instead, by specifying the field on a surface for $t > 0$, is equivalent to impose a time varying boundary condition with zero initial data. This way of proceeding produces the correct primary field without altering the well posedness of the PDE problem because:

- if the space-time dependence in (7.1) is used, the data for the primary field are specified on the correct characteristics curves of the PDE system;
- the boundary conditions for the primary field are specified in a “soft” way, i.e. using the superposition principle, not by fixing the values in (7.1) at the surface locations (see below for a more detailed explanation). In this way the secondary field is free to pass through this boundary surface without conflicts with its characteristic directions which are not known and depend on the structure of the target.

A two dimensional schematic view of the plane wave “injection” is given in Fig. 7.1:

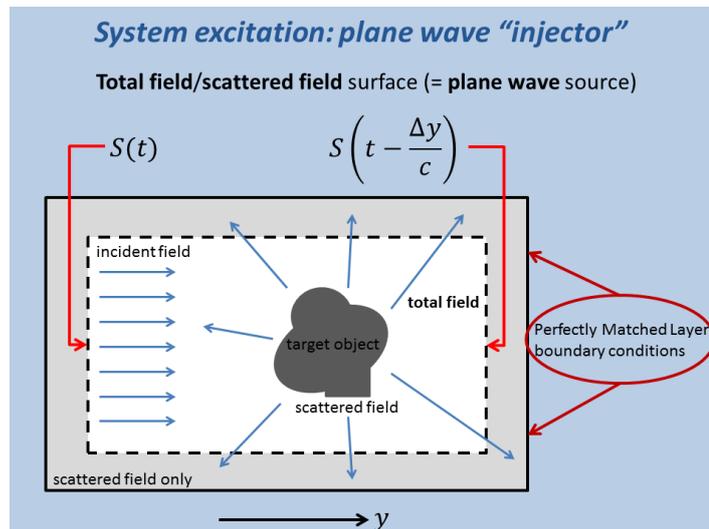


Figure 7.1: Two dimensional view of the FDTD system excitation with an impinging light beam.

Including the impinging light beam in the manner just described, produces

two regions in the spatial computational domain spanning N_x , N_y and N_z Yee cells along the x , y and z axes respectively: an inner region in which the total field (TF) is found (a superposition of the incident and scattered fields, in a mix which however cannot be resolved in the two single constituents), and an outer shell (the light grey area in Fig. 7.1) in which only the scattered field (SF) is present. This two regions have a surface in common referred to as \mathcal{S} below: it is represented by the dashed rectangle in Fig. 7.1. At the outer surface \mathcal{S}' of this latter region of space (the solid line rectangle in Fig. 7.1) begins, not shown in the figure, the region corresponding to the CPML boundary conditions, used to simulate an infinite spatial extent around the target. The CPML region constitutes an extra outer shell extending further the spatial computational domain comprising the $N_x N_y N_z$ Yee cells.

7.1 Propagation along a coordinate axis

If the y coordinate axis is chosen as the propagation direction of the initial impinging beam, then $\hat{u} = \hat{y}$ in (7.1) and the concerned independent variables are t and y . The TF/SF surface \mathcal{S} is chosen to be inflated with respect to \mathcal{S}' of an equal number of Yee cells on both sides, in such a way to make space for the scattered field (SF) region leaving centered the total field (TF) one. The “soft” way of including the excitation consists in:

- correcting, after each updating FDTD iteration, all the \vec{E} tangential fields *on* \mathcal{S} , which are considered to be total field, by *adding* to them a contribution as if, to the relevant \vec{H} field components in the finite-difference equations (5.2), the incident term (7.1b) has been added a posteriori.
- correcting, after each updating FDTD iteration, all the \vec{H} tangential fields *just outside* \mathcal{S} by half cell edge size, which are considered to be scattered field, by *subtracting* to them a contribution as if, to the relevant \vec{E} field components in the finite-difference equations (5.1), the incident term (7.1a) has been subtracted a posteriori.

From a programming code point of view, the corrections are made for all the tangential field components on, or just outside, \mathcal{S} at each FDTD iteration, but the propagation delay in the argument of (7.1) determines if and where the correction term will be effectively included, and in a coherent way with the propagation of the traveling exciting signal.

7.2 Propagation along arbitrary directions

If \hat{u} in (7.1) is an arbitrary unit vector, the procedure for including the linearly polarized plane wave is substantially the same as above, with the difference that is less a priori evident which are the tangential field components affected by the correction and which are the cartesian components to which add/subtract the incident field contributions (7.1). It is convenient to introduce a rotated orthogonal reference frame for \hat{u} , \vec{E} and \vec{H} and then, after transforming to the fixed reference frame of the spatial computational grid, to calculate scalar products to identify which are the components on \mathcal{S} to be affected by the corrections and the amplitude of the correction, which depends on the projection on \mathcal{S} itself. This procedure is computationally more cumbersome than that for the propagation along a coordinate axis, because it requires a larger number of mathematical operations to be performed by the code. Again, as in the preceding case, the propagation delay in the argument of (7.1) determines if and where the correction term will be effectively included, and in a coherent way with the propagation of the traveling exciting signal. In this case the equal phase planes are in general not parallel to the coordinate planes.

Chapter 8

Algorithm modification for dispersive materials

The present Chapter concerns with the possibility of including, in the FDTD algorithm, material media exhibiting temporal dispersion. This means a temporal non-locality between \vec{D} and \vec{E} , as previously described in Section 1.2 of Chapter 1, resulting in a still linear one: a time convolution, but no more simply multiplicative relation between the two fields in the time domain — it will be simply multiplicative, in general, in the frequency domain — and in the complex electric permittivity ϵ dependence on the (angular) frequency. The importance of achieving this possibility is connected to the multifrequency analysis with a *single* FDTD run described previously in Section 5.3: in the time domain one cannot work with $\epsilon(\omega)$ in a tabular form and only if she/he is able to correctly track this temporal non-locality it will make sense the update and memorization, at every FDTD time iteration, of a Discrete Fourier Transform (DFT) variable at several different values of ω . The details of the generalized macroscopic linear relation between \vec{D} and \vec{E} , the kernel of the time convolution, will depend on the physical polarization mechanism assumed to hold for a given material.

8.1 Recursive convolution

This approach to the FDTD description of dispersive materials is based on a representation of the integral in the time convolution relating \vec{D} and \vec{E} [...] as a discrete sum, consistently with the temporal sampling at a time step δt and once the specific functional form of the kernels has been calculated. In [37,38] it is demonstrated that for these kernels, corresponding [...], the time convolution can be updated without memorizing the whole field evolution

history from the beginning of the simulation, but that it can be *recursively updated* with a minimal amount of supplementary memory. Because this recursive convolution method is very similar to the one used to update the absorbing layer in the CPML boundary conditions described in Chapter 6, it will not be described further here.

8.2 Auxiliary differential equation

Another approach to the FDTD description of dispersive media is by means of a finite difference version of a suitable auxiliary differential equation, coupled with the usual finite difference expressions of the Maxwell's curl equations. This has to be done at every spatial location for the electric field components of the Yee grid belonging to a dispersive material. From a fairly general form for the relative complex electric permittivity in the (angular) frequency domain (for a time-harmonic factor $e^{-i\omega t}$) [45, 46]:

$$\epsilon_r(\omega) = \epsilon_{r,\infty} + \sum_{\ell=1}^N \frac{a_{\ell,0} - i a_{\ell,1} \omega}{b_{\ell,0} - i b_{\ell,1} \omega - b_{\ell,2} \omega^2} = \epsilon_{r,\infty} + \chi(\omega) \quad (8.1)$$

(i is the imaginary unit) where contributions from N terms have been considered and $\epsilon_{r,\infty}$ is from a Dirac delta contribution in the time convolution, i.e. an instantaneous response of the material. It could be referred to as a “static” contribution if one means it is the only contribution when dispersion is completely neglected by discarding the sum term in (8.1). Otherwise it actually corresponds to the $\omega \rightarrow \infty$ limit of $\epsilon_r(\omega)$, which is not the static limit. The real parameters $a_{\ell,0}, a_{\ell,1}, b_{\ell,0}, b_{\ell,1}, b_{\ell,2}$ for $\ell = 1, \dots, N$ are chosen to fit the experimental tabulated data in the frequency range of interest, or they can be determined theoretically on the base of a model for the material response. They should also properly normalized in such a way that:

$$\sum_{\ell=1}^N \frac{a_{\ell,0}}{b_{\ell,0}} = \epsilon_{r,s} - \epsilon_{r,\infty}$$

where $\epsilon_{r,s}$ is the true static ($\omega = 0$) relative electric permittivity.

Expression (8.1) encompasses several polarization behavior models (see Section 1.2):

- Debye: by setting $a_{\ell,1} = b_{\ell,2} = 0$, for $\ell = 1, \dots, N$;
 - Drude: by setting $a_{\ell,1} = b_{\ell,0} = 0$, for $\ell = 1, \dots, N$;
 - Lorentz: by setting $a_{\ell,1} = 0$, for $\ell = 1, \dots, N$;
 - Critical Points: as such, (8.1) corresponds to an N Critical Points model.
- From the general relation, in which usually $\epsilon_{r,\infty} = 1$, see (1.8):

$$\vec{D} = \epsilon_o \epsilon_{r,\infty} \vec{E} + \vec{P},$$

and the assumed linear multiplicative constitutive relation between the polarization vector \vec{P} (whose units are *Coulomb/m²* or dipole moment per unit volume) and \vec{E} , through the *electric susceptibility* $\chi(\omega)$, remembering that, in general:

$$\vec{\nabla} \cdot \vec{P} = -\rho_{\text{pol}}$$

where ρ_{pol} is the electric volume density of the *polarization* charge, and considering its continuity equation (in the frequency domain with the correspondence $-i\omega \Leftrightarrow \frac{\partial}{\partial t}$):

$$\vec{\nabla} \cdot \vec{j}_{\text{pol}} - i\omega \rho_{\text{pol}} = 0$$

where \vec{j}_{pol} is the polarization current density vector, one has:

$$\vec{j}_{\text{pol}} = -i\omega \vec{P} = -i\omega \chi(\omega) \vec{E}$$

i.e., the time derivative of \vec{P} . If the above definition for \vec{D} is time derived to use it in the right hand side of (1.1d) Maxwell's curl equation, one obtains (in the time domain):

$$\vec{\nabla} \times \vec{H} = \epsilon_o \epsilon_{r,\infty} \frac{\partial \vec{E}}{\partial t} + \sigma \vec{E} + \sum_{\ell=1}^N \vec{j}_{\ell} \quad (8.2)$$

where the last sum term on the right represents precisely \vec{j}_{pol} (in the time domain). Also, in the last equation the static conductivity term has been retained, but the free current source term has been omitted (external free sources are not included inside dispersive media). Now, in the frequency domain:

$$\vec{j}_{\ell} = -i\omega \epsilon_o \frac{a_{\ell,0} - i a_{\ell,1} \omega}{b_{\ell,0} - i b_{\ell,1} \omega - b_{\ell,2} \omega^2} \vec{E}$$

($\ell = 1, \dots, N$) and there is the need to get a time domain expression for \vec{j}_{ℓ} , or at least for its sampled values, to be used in the discretized version of (8.2). To this end, the above expression is simply rewritten, after multiplication of both members by the denominator in the right one, as:

$$b_{\ell,0} + b_{\ell,1}(-i\omega) \vec{j}_{\ell} + b_{\ell,2}(-i\omega)^2 \vec{j}_{\ell} = a_{\ell,0} \epsilon_o (-i\omega) \vec{E} + a_{\ell,1} \epsilon_o (-i\omega)^2 \vec{E}$$

($\ell = 1, \dots, N$) and, with the usual identification $-\omega \Leftrightarrow \frac{\partial}{\partial t}$, which also implies $-\omega^2 \Leftrightarrow \frac{\partial^2}{\partial t^2}$, this becomes the *auxiliary differential equation*:

$$b_{\ell,0} \vec{j}_\ell + b_{\ell,1} \frac{\partial \vec{j}_\ell}{\partial t} + b_{\ell,2} \frac{\partial^2 \vec{j}_\ell}{\partial t^2} = a_{\ell,0} \epsilon_o \frac{\partial \vec{E}}{\partial t} + a_{\ell,1} \epsilon_o \frac{\partial^2 \vec{E}}{\partial t^2}$$

$\ell = 1, \dots, N$. By spatially and temporally sampling \vec{j}_ℓ like \vec{E} and approximating the first time derivatives as:

$$\frac{\partial \vec{j}_\ell}{\partial t} \approx \frac{\vec{j}_\ell^{n+1} - \vec{j}_\ell^{n-1}}{2\delta_t}$$

$$\frac{\partial \vec{E}}{\partial t} \approx \frac{\vec{E}^{n+1} - \vec{E}^{n-1}}{2\delta_t},$$

and the second time derivatives as:

$$\frac{\partial^2 \vec{j}_\ell}{\partial t^2} \approx \frac{\vec{j}_\ell^{n+1} - 2\vec{j}_\ell^n + \vec{j}_\ell^{n-1}}{\delta_t^2}$$

$$\frac{\partial^2 \vec{E}}{\partial t^2} \approx \frac{\vec{E}^{n+1} - 2\vec{E}^n + \vec{E}^{n-1}}{\delta_t^2},$$

inserting in the above auxiliary differential equation and solving for \vec{j}_ℓ^{n+1} , one obtains:

$$\boxed{\vec{j}_\ell^{n+1} = \alpha_\ell \vec{j}_\ell^n + \beta_\ell \vec{j}_\ell^{n-1} + P_\ell \vec{E}^{n+1} + Q_\ell \vec{E}^n + R_\ell \vec{E}^{n-1}} \quad (8.3)$$

where:

$$\alpha_\ell = \frac{b_{\ell,2} - b_{\ell,0} \delta_t^2}{b_{\ell,2} + b_{\ell,1} \delta_t} \quad (8.4a)$$

$$\beta_\ell = \frac{b_{\ell,1} \delta_t - b_{\ell,2}}{b_{\ell,2} + b_{\ell,1} \delta_t} \quad (8.4b)$$

$$P_\ell = \frac{a_{\ell,0} \delta_t + 2a_{\ell,1} \epsilon_o}{b_{\ell,2} + b_{\ell,1} \delta_t} \quad (8.4c)$$

$$Q_\ell = -\frac{4a_{\ell,1}}{b_{\ell,2} + b_{\ell,1} \delta_t} \epsilon_o \quad (8.4d)$$

$$R_\ell = -\frac{2a_{\ell,1} - a_{\ell,0} \delta_t}{b_{\ell,2} + b_{\ell,1} \delta_t} \epsilon_o. \quad (8.4e)$$

Discretizing (8.2) gives:

$$\left[\vec{\nabla} \times \vec{H}\right]^{n+\frac{1}{2}} = \epsilon_o \epsilon_{r,\infty} \frac{\vec{E}^{n+1} - \vec{E}^n}{\delta_t} + \sigma \frac{\vec{E}^n + \vec{E}^{n+1}}{2} + \sum_{\ell=1}^N \vec{j}_\ell^{n+\frac{1}{2}}$$

in which $\vec{j}_\ell^{n+\frac{1}{2}}$ is calculated by inserting the time average of (8.3):

$$\vec{j}_\ell^{n+\frac{1}{2}} = \frac{\vec{j}_\ell^n + \vec{j}_\ell^{n+1}}{2} = \frac{1 + \alpha_\ell}{2} \vec{j}_\ell^n + \frac{\beta_\ell}{2} \vec{j}_\ell^{n-1} + \frac{P_\ell}{2} \vec{E}^{n+1} + \frac{Q_\ell}{2} \vec{E}^n + \frac{R_\ell}{2} \vec{E}^{n-1},$$

from which the updated value of \vec{E} is:

$$\boxed{\begin{aligned} \vec{E}^{n+1} &= C_1 \vec{E}^n + C_2 \vec{E}^{n-1} \\ &+ C_3 \left\{ \left[\vec{\nabla} \times \vec{H}\right]^{n+\frac{1}{2}} + \frac{1}{2} \sum_{\ell=1}^N \left[(1 + \alpha_\ell) \vec{j}_\ell^n + \beta_\ell \vec{j}_\ell^{n-1} \right] \right\} \end{aligned}} \quad (8.5)$$

where:

$$C_1 = \frac{2\epsilon_o \epsilon_{r,\infty} - \delta_t \left(\sigma + \sum_{\ell=1}^N Q_\ell \right)}{2\epsilon_o \epsilon_{r,\infty} + \delta_t \left(\sigma + \sum_{\ell=1}^N P_\ell \right)} \quad (8.6a)$$

$$C_2 = -\frac{\delta_t \sum_{\ell=1}^N R_\ell}{2\epsilon_o \epsilon_{r,\infty} + \delta_t \left(\sigma + \sum_{\ell=1}^N P_\ell \right)} \quad (8.6b)$$

$$C_3 = -\frac{2\delta_t}{2\epsilon_o \epsilon_{r,\infty} + \delta_t \left(\sigma + \sum_{\ell=1}^N P_\ell \right)}. \quad (8.6c)$$

Thus for dispersive media the FDTD algorithm is as follows [45]:

- update \vec{E}^{n+1} using (8.5) with (8.6), from \vec{E}^n , $\vec{H}^{n+\frac{1}{2}}$, \vec{j}_ℓ^n and \vec{j}_ℓ^{n-1} ($\ell = 1, \dots, N$);
- update \vec{j}_ℓ^{n+1} ($\ell = 1, \dots, N$) using (8.3) with (8.4), using also the \vec{E}^{n+1} just computed;
- update $\vec{H}^{n+\frac{3}{2}}$ the standard FDTD algorithm (5.1) using $\vec{H}^{n+\frac{1}{2}}$ and the \vec{E}^{n+1} just computed.

The standard FDTD algorithm modified in such a way to include temporal dispersive materials requires extra memory allocation to store the \vec{j}_ℓ^n , \vec{j}_ℓ^{n-1} ($\ell = 1, \dots, N$) and \vec{E}^{n-1} components on the spatial grid. Moreover, a larger

number of floating point operations must be performed, for each FDTD time iteration, to get the updated electric field values, which implies longer simulation running times. However, the modified FDTD algorithm is restricted to the spatial grid points of the whole computational volume corresponding to the regions occupied by dispersive media. The multiterm ℓ index variability in (8.1) makes possible to use different polarization models simultaneously for a given material, to better fit the dispersion curve of its complex electric permittivity in the frequency range of interest. In [46] is used an approach substantially equivalent to the one just described, but using directly the polarization vectors \vec{P}_ℓ in place of the polarization currents. In [47] are deduced exact stability criteria for these algorithms.

Chapter 9

FDTD algorithm parallelization

When the system, the interaction with light of which has to be studied, is a large one or the level of details required asks for very small δ_x , δ_y , δ_z space steps sizes, the computations will need huge amounts of memory (RAM), which increases linearly with the total number of Yee cells allocated: $N_{cells} = N_x \times N_y \times N_z$, according to (1 byte = 8 bit):

$$\mathcal{N}_{bytes} = N_{cells} [4 + (1 + n_{freqs})6\eta] \quad (9.1)$$

where n_{freqs} is the number of frequencies one wants to analyze in a single run of the FDTD code (as previously described in a devoted Section about frequency domain analysis in Chapter 5) and $\eta = 4$ or 8, depending on the choice for the floating point variables precision, respectively “single” or “double”. Double precision is generally used, to minimize round-off error accumulation during the FDTD iterations cycles. Also, the execution time of the FDTD code increases linearly with N_{cells} , because the number of operations in the code scales $\propto N_{cells}$. When the computational burden, i.e. the memory request and the execution time, becomes exceedingly high, one has to resort to High Performance Computing (HPC) machines, like the one depicted in Fig. 9.1:



Figure 9.1: An HPC machine for large parallel computations.

which make allowable and can handle all the memory required. Basically

they are an ensemble of multicore CPUs (Central Processing Units), interconnected through an high speed dedicated network for data exchange among them: the multicomputer MIMD (Multiple Instruction Multiple Data) architecture. Each CPU is mounted on a board with a given amount of RAM memory. The memory in a board can be accessed by its CPUs more quickly than that on different boards, which instead has to be accessed through the communication network, whose speed is a critical factor. Many of the calculations for the present work have been executed on a IBM-BlueGene with a total of 163,840 cores. To take advantage of such computational capabilities, *parallel computation* has to be used, which means that the instructions and data constituting the code have to be suitably rearranged in such a way that many instructions are executed as concurrently, but in general not independently, *processes* by the different CPU cores on smaller parts of the data — in the present case the \vec{E} and \vec{H} on the spatial grid. This implies also a decrease in the execution time of the overall code. Theoretically, the smaller the parts of data processed by each core and the shorter should be the execution times, with a linear scaling. This is not exactly true, however, because the communication through the connecting network has its own costs in terms of latency and transmission time. Thus there is an optimal size of the data subdivision among different cores which is the most effective for computational speed-up. One of the most powerful tools for code parallelization is by using the Message Passage Interface (MPI) [10–12], a library of routines implementing various functionalities for point-to-point or collective communication among processes. Each process corresponds, in the MPI paradigm, to a subdivision unit of the whole data set and the correlated subset of code instructions, and runs on a distinct physical CPU core.

9.1 Domain decomposition

The parallelization strategy for the FDTD code — i.e., the task and data subdivision of the whole algorithm described in the preceding Chapters — adopted in the present work, consists in the decomposition of the whole spatial grid, including the outer PML shell for the boundary truncation, into disjoint subdomains (blocks of data) as suggested by Fig. 9.2 where the subdomains are parallelepipeds due to the rectangular cartesian coordinates used to formulate the algorithm itself. A partial accomplishment of this task has been already obtained by the author of the present work, and coauthors, in [31]. Each block comprises many Yee cells (like the ones in Figs 5.1 and 5.2) and the FDTD iterations for each such subdomain are executed on a distinct CPU core. The single subdomain is uniquely identified with

an ordering integer number and with the same number will be identified the process executing it. A subdomain can also be alternatively identified with a triple of integers giving its position with respect to the cartesian axes. Between the two numbering systems there is a one-to-one correspondence. The whole algorithm proceeds with many processes allocating the memory for \vec{E} and \vec{H} for their own part of the total grid. The processes then run concurrently each one performing the cycle of FDTD iterations for the time evolution of the fields inside their own part of the spatial grid (proper computation). Although concurrently, the processes cannot run independently, because a synchronization among them is necessary for a coherent and consistent solution for \vec{E} and \vec{H} on the whole grid. Such a synchronization is obtained by exchanging, among the various processes, the tangential \vec{E} and \vec{H} components to the exterior surface of the subdomains, at the beginning of each FDTD iteration (communication). By fixing, taken from the outside (ultimately, from the exterior PML shell), the tangential components of the electric and magnetic fields on the outer surface of a grid block guarantees, at each time step, the univocity of the solution inside it. Obviously one process needs to exchange data only with the processes executing blocks which are, topologically, first neighbors of the one considered. See Fig. 9.3. Such a way of operating the parallelization is affordable due the scaling law $\propto \sqrt{N}$ of the data on a surface (communication), with respect to those in a bulk (proper calculation), which scale $\propto N$ (in N is the number of cells in the subdomain). In any case a high speed connecting network is required for good performances. Fig. 9.4 gives a schematic example of the exchange of tangential \vec{E} and \vec{H} data on an outer surface in common to two neighboring subdomains running on two distinct MPI processes (CPU cores).

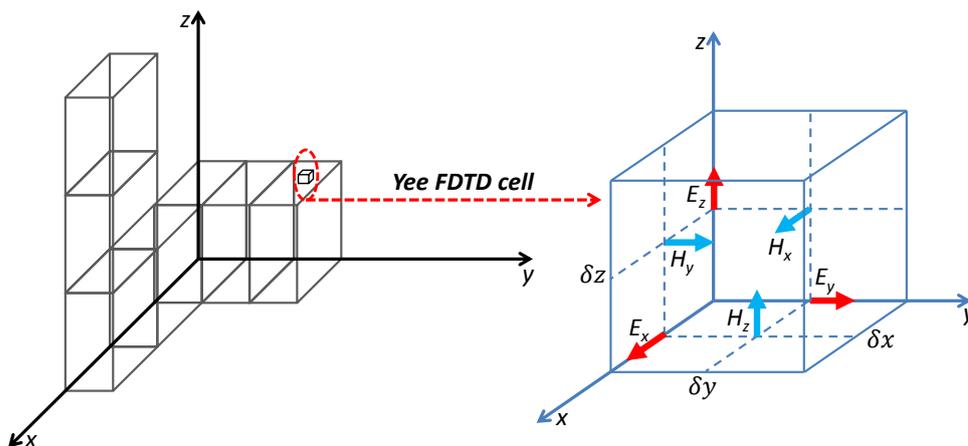


Figure 9.2: Domain decomposition into subdomains (blocks), of the whole FDTD spatial grid for parallel computation of the algorithm. Each block comprises many Yee cells.

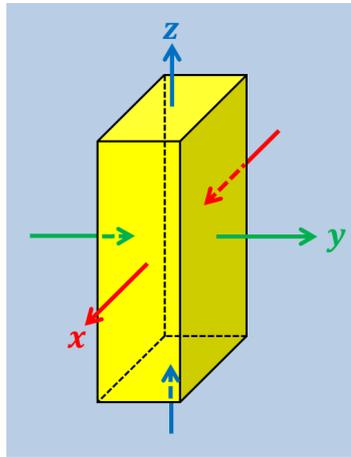


Figure 9.3: In FDTD parallel computation, a subdomain (yellow) process needs to communicate with its six first neighbors (see text).

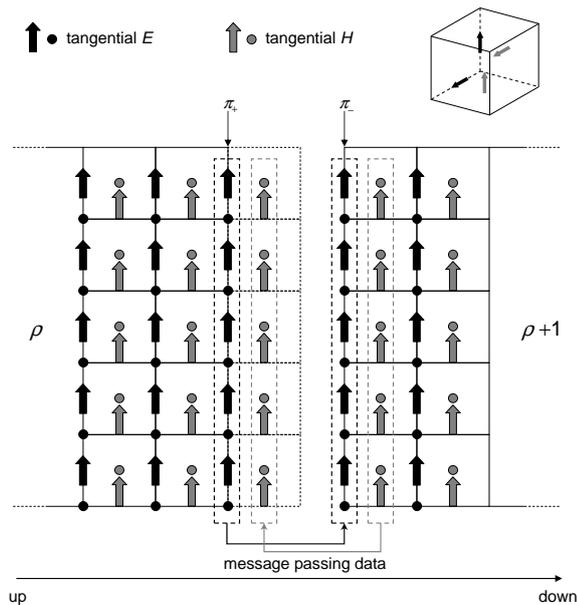


Figure 9.4: Scheme in one dimension of the exchange of the electric and magnetic tangential data between two neighboring processes identified by the two consecutive numbers ρ and $\rho + 1$ (see text).

These are identified with the two consecutive integer numbers ρ and $\rho + 1$. The scheme shown is one dimensional, describing the exchange along a single coordinate axis direction. It should be thought as applied to all the six faces of the block in Fig. 9.3. It should also be noted that the electric and magnetic tangential components are exchanged in opposite directions with respect to the process ordering along a given coordinate axis. This is required

by the bulk FDTD algorithm, because to calculate the outer tangential \vec{E} components of ρ one needs the tangential \vec{H} components which are half space step further outside, and thus come from the neighbor process $\rho + 1$. To calculate instead the tangential \vec{H} components which are half space step inward from the outer surface of $\rho+1$, one needs the tangential \vec{E} components which are exactly on its outer surface and which come from ρ . The data point-to-point exchange operations like this one are best performed by means of the MPI function `Sendrecv(...)` [10–12], which allows the correct matching of the sending and receiving instruction pairs on the two sides. If a `Send(...)` MPI command is not matched to a corresponding `Recv(...)` command and vice versa, the program will stall without producing any results. This is a common deadlock occurring in parallel programming which one must be aware of. Another bottleneck, potentially able to greatly slow down the execution of the parallel FDTD code is when the amounts of computation of the various processes are unbalanced, typically because the whole spatial grid is not evenly decomposed into subdomains. If this is the case, it becomes hard to get any synchronization of the processes at any FDTD iteration, because inherently some of them will be occupied in calculations while others would be ready for the communication phase, having finished the computations inside their subdomains. To avoid such an irregular pacing of the code, a good balancing of the amount of work among processors has to be preventively assured.

9.2 Suitable MPI data structures

For a correct and efficient MPI parallel programming of the FDTD algorithm, great importance have the data structures, allocated by each process to contain the electric and magnetic field components values of the corresponding subdomain, and the way in which they are addressed on the outer subdomain surface when they are exchanged among topologically contiguous processes. A particular concern is that of allocating three-dimensional arrays with the assurance that all the elements occupy *a single contiguous memory area*. In this way slices can be extracted from the array perpendicularly to each one of the three mutually orthogonal axes according to a predefined pattern and safely communicated to different processes using the previously mentioned MPI routines. These routines require a starting data address and a buffer size and necessarily must operate on data occupying contiguous memory locations or extracted from a contiguous memory area according to the predefined pattern. Thus, in the present work, a three-dimensional array `p[·][·][·]` is allocated with the following instructions (C++ code):

```

int dim2 = ny*nz;

int dim3 = nx*ny*nz;

double ***p;

p = new double** [nx];

for (i = 0; i < nx; ++i) p[i] = new double* [ny];

p[0][0] = new double [dim3];    // (*)

for (i = 0; i < nx; ++i)
{
    shift1 = i*dim2;

    for (j = 0; j < ny; ++j)
    {
        shift2 = j*nz;

        if (i != 0 || j != 0) p[i][j] = p[0][0] + shift1 + shift2;
    }
}

```

where n_x , n_y , n_z are the sizes of the three-dimensional array. With this dynamic allocation, the data are stored ordering them along the k -axis first, then along the j -axis and finally along the i -axis. The “linear” indexing being:

$$\ell = (i - 1)n_y n_z + (j - 1)n_z + k$$

with $i = 1, \dots, n_x$, $j = 1, \dots, n_y$, $k = 1, \dots, n_z$ and $\ell = 1, \dots, n_x n_y n_z$. The memory area containing all the $n_x n_y n_z$ stored values starts at the address of $p[0][0][0]$ (see the code line marked with $(*)$): the **new** command guarantees a single contiguous memory area, but multiple call of the **new** command could generate multiple areas which are not contiguous and do not form a single uncut memory segment). The other **new** calls in the above code sample, simply allocate pointers which are then shifted to point to the appropriate address, starting from $p[0][0][0]$, according to the chosen data ordering (last line of code).

Once a suitable “container” for the field component values of a spatial grid subdomain has been defined according to the above structure, ij -, ik -, jk -

slices can easily be extracted from it by defining extra MPI datatypes using the `Create_vector(...)` MPI routine [10–12], with the appropriate `count`, `blocklength` and `stride` parameters, which specify the regular pattern on the base of which to select the data from the array.

- For an `ij`-slice one has:
`count = nx*ny`, `blocklength = 1` and `stride = nz`.
- For an `ik`-slice one has:
`count = nx`, `blocklength = nz` and `stride = ny*nz`.
- For a `jk`-slice one has:
`count = ny*nz`, `blocklength = 1` and `stride = 1`.

The meaning of the parameters is as follows: `blocklength` specifies how many old datatypes (here `double`) in the array the new datatype contains; `count` specifies how many pieces of length `blocklength` the new datatype contains; `stride` specifies the “distance”, in old datatype elements, between two consecutive elements in the new datatype. For example, a `stride = 1` means that in the array the data are consecutive. Using these new datatypes in the `Sendrecv(...)` MPI functions it is an easy matter to communicate data among processes in an efficient and ordered way, thus minimizing the latency time costs caused by multiple unnecessary calls to them.

Part III

Applications of the FDTD method

Chapter 10

Validation of the implemented FDTD method

10.1 Field inside and scattered off a sphere

To validate the in-house computer code produced for the present work, and implementing the FDTD algorithm as described in the previous Part II, a comparison test is made here between the results obtained numerically for a sphere of given radius impinged by an electromagnetic, linearly polarized plane wave and the analytical solution for the same system, as given in Section 1.3 of Part I. Even the analytical solution has to be evaluated numerically, after implementation of the physico-mathematical special functions required for the solution. Because such analytical solution is in the form of a series expansion, many terms of this are calculated and summed up, to ensure good convergence of the data serving as the comparison reference. More than 30 terms of the expansion have been considered.

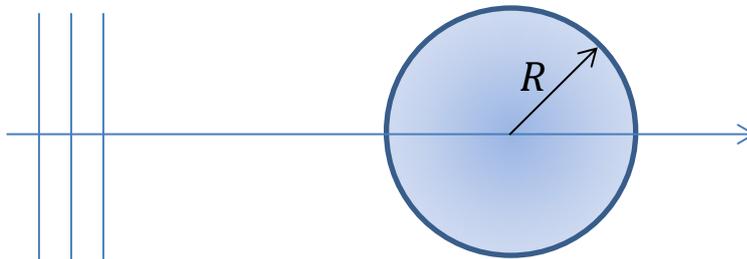


Figure 10.1: 2D scheme of a linearly polarized e.m. traveling plane wave impinging on a sphere.

According to the scheme of Fig. 10.1, a dielectric sphere of relative electric

permittivity $\varepsilon_r = 4$ and radius $R = 4.5$ cm is considered in the electromagnetic field of a linearly polarized monochromatic plane wave of frequency $\nu = 2.5$ GHz (corresponding to a vacuum wavelength $\lambda \cong 12$ cm), impinging on it. The resulting FDTD and analytical field distributions of the field component in the polarization direction (normalized to the incident one), are shown in Fig. 10.2 along the propagation axis, passing through the sphere's center, which is located at 26 cm.

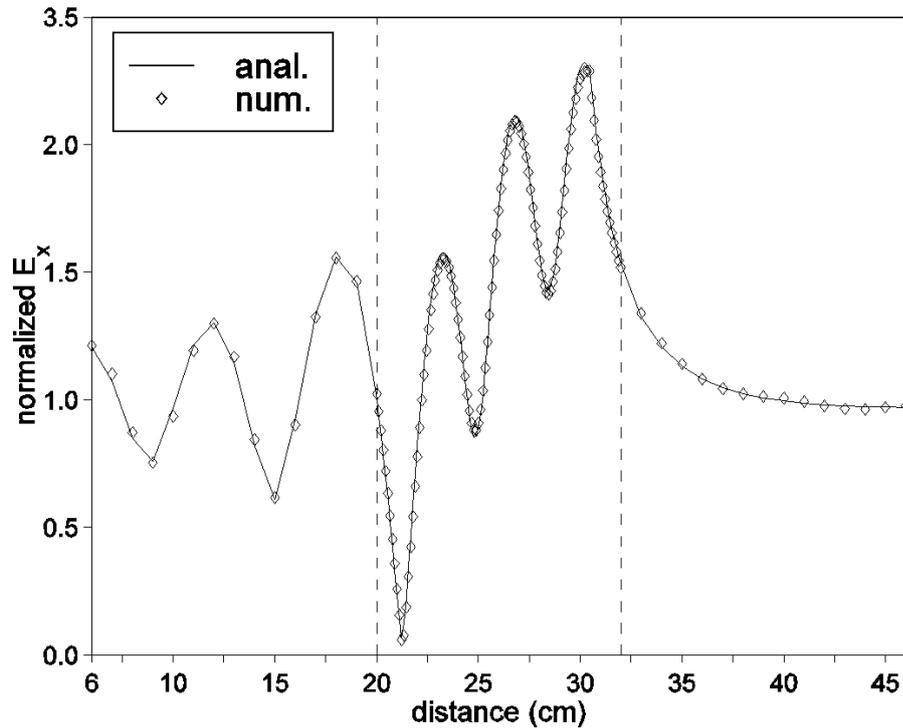


Figure 10.2: Compared field distributions along the sphere's axis. The sphere's center is at 26 cm (see text).

in which it is apparent a fairly good agreement. Another comparison test is made on a gold sphere with radius $R = 96$ nm for which the spectral extinction and absorption coefficients are calculated. The *absorption* and *scattering coefficients* are here defined, at any given frequency (or vacuum wavelength) as the power absorbed (and dissipated into heat) or scattered in all directions (i.e., in the whole solid angle around it) by the sphere, normalized to the incident power. This latter is the intensity of the incident plane wave ($Watt/m^2$) times the geometrical cross section $A = \pi R^2$ of the sphere itself. The *extinction coefficient* is the sum of the scattering and absorption coefficients. Practically, these coefficients are pure numbers expressing the

effective cross section measured in units of the geometric one, for scattering or absorption. The comparison graphs are shown in Fig. 10.3:

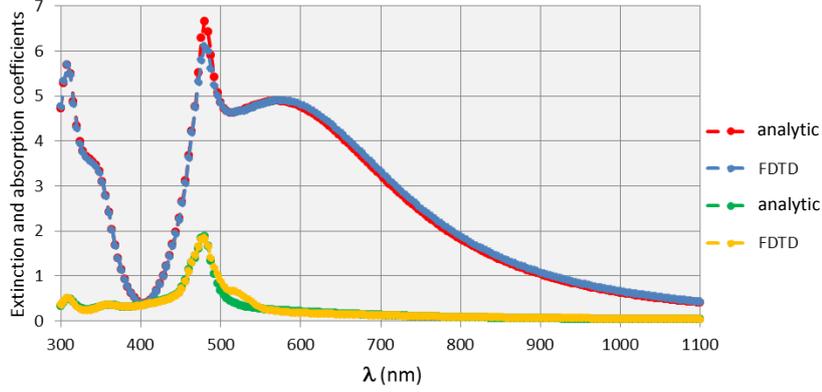


Figure 10.3: Extinction and absorption coefficients for an Au sphere with radius $R = 96$ nm.

which shows a good agreement between the expected and numerically calculated values over the $300 \div 1100$ vacuum wavelength range. The scattering and extinction coefficients Q_{scatt} and Q_{ext} respectively, can be directly calculated analytically by means of the solution expansion coefficients (1.15a) and (1.15b), given previously in Section 1.3, according to [5]:

$$Q_{scatt}(\lambda) = \frac{1}{2} \left(\frac{\lambda}{\pi R} \right)^2 \sum_{m=1}^{\infty} (2m+1) [|a_m^r(\lambda)|^2 + |b_m^r(\lambda)|^2] \quad (10.1a)$$

$$Q_{ext}(\lambda) = \frac{1}{2} \left(\frac{\lambda}{\pi R} \right)^2 \operatorname{Re} \left\{ \sum_{m=1}^{\infty} (2m+1) [a_m^r(\lambda) + b_m^r(\lambda)] \right\} \quad (10.1b)$$

at any given frequency. The corresponding $Q_{scatt}^{num.}$ and $Q_{abs}^{num.} = Q_{ext}^{num.} - Q_{scatt}^{num.}$ values, as obtained from the numerical FDTD simulation, and compared with the “exact” ones in Fig. 10.3, are evaluated through the integration of the *real part* of the complex Poynting vector on a parallelepiped surface completely enclosing the metallic sphere. For $Q_{abs}^{num.}$ the integration surface is wholly inside the total field region (see Fig. 7.1). For $Q_{scatt}^{num.}$ the parallelepiped surface is contained in the outer shell in which only the scattered field is calculated. For this evaluation it is important the “injection mechanism” for the plane wave excitation described in Chapter 7. The scattered near field can be transformed in the true scattered radiation field by means of the Kirchhoff integral formula (2.16) described early. The power fluxes obtained

by the two field distributions coincide, because reactive components in the scattered near field do not contribute to the real part of the Poynting surface integral. However, equation (2.16) allows the detailed calculation of the scattered radiation amplitude at any direction of the solid angle around the sphere and could equally well be used for an arbitrary shaped scatterer.

The FDTD calculations for the gold sphere have been performed implementing the algorithm for dispersive materials described in Chapter 8 and by using a Drude-Sommerfeld + double Critical Points model as given early in equations (1.7), (1.10) and (1.13). The parameters used inside those expressions are found in [48]:

$$\begin{aligned}
 \varepsilon_{r,\infty} &= 1.1431 \\
 \omega_{\text{pl}} &= 1.3202 \times 10^{16} \text{ rad/sec} \\
 \gamma &= 1.0805 \times 10^{16} \text{ rad/sec} \\
 A_1 &= 0.26698 \\
 \phi_1 &= -1.2371 \\
 \Omega_1 &= 3.8711 \times 10^{15} \text{ rad/sec} \\
 \Gamma_1 &= 4.4642 \times 10^{14} \text{ rad/sec} \\
 A_2 &= 3.0834 \\
 \phi_2 &= -1.0968 \\
 \Omega_2 &= 4.1684 \times 10^{15} \text{ rad/sec} \\
 \Gamma_2 &= 2.3555 \times 10^{15} \text{ rad/sec} .
 \end{aligned}$$

For silver they are instead [48]:

$$\begin{aligned}
 \varepsilon_{r,\infty} &= 15.833 \\
 \omega_{\text{pl}} &= 1.3861 \times 10^{16} \text{ rad/sec} \\
 \gamma &= 4.5841 \times 10^{13} \text{ rad/sec} \\
 A_1 &= 1.0171 \\
 \phi_1 &= -0.93935 \\
 \Omega_1 &= 6.6327 \times 10^{15} \text{ rad/sec} \\
 \Gamma_1 &= 1.6666 \times 10^{15} \text{ rad/sec} \\
 A_2 &= 15.797 \\
 \phi_2 &= 1.8087 \\
 \Omega_2 &= 9.2726 \times 10^{17} \text{ rad/sec} \\
 \Gamma_2 &= 2.3716 \times 10^{17} \text{ rad/sec}
 \end{aligned}$$

and are used for the simulation from which Fig. 10.4 is extracted. It is a snapshot from the time evolution of a plane wave pulse, like the one represented by the solid line in Fig. 5.6, impinging on an Ag sphere with $R = 60$ nm. The color map represents the module of the resulting electric field with the pulse, propagating from left to right, just past the sphere. On this latter it is apparent a dipolar charge separation on the “downwind” hemisphere.

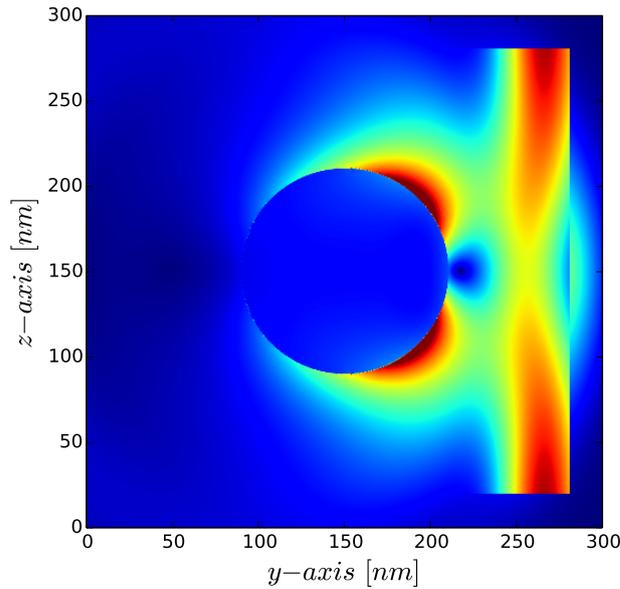


Figure 10.4: Color map snapshot of the resulting electric field after a plane wave pulse has impinged, from left to right, on an Ag sphere of radius $R = 60$ nm. The pulse is just past the sphere.

The simulation in Fig. 10.4 has been performed with Yee cell step sizes $\delta_x = \delta_y = \delta_z = 0.25$ nm, with a computational volume with an overall number of $1200 \times 1200 \times 1200$ cells.

10.2 Scalability of the parallelized code

An important aspect of the code parallelization, as has been explained in Chapter 9, and particularly in Section 9.1, is its ability to maintain constant performances for an increasingly finer decomposition of the spatial computational volume. This means an increasing speed-up (i.e., a lesser execution time) with the number of MPI processes launched for the code execution (which is paid with a greater computational resources allocation, i.e., constant performances). However, a finely grained decomposition, the so called granularity, implies a lot of MPI process communication, which take place at

the expense of the bulk calculations inside any given spatial subdomain. This penalizes the overall speed-up and constitutes an upper saturation limit preventing from an ideal linear increase. Fig. 10.5, which reports the scalability analysis of the FDTD code for the present work, as applied to the electromagnetic field calculation for the metal sphere of the previous Section, illustrates this behavior. The code runs on a BlueGene/Q supercomputer, where each node comprises 16 cores, in a 1.6 GHz PowerPC based CPU (PowerPC A2), and has a 16 GB of shared memory. The total node number is 2048.

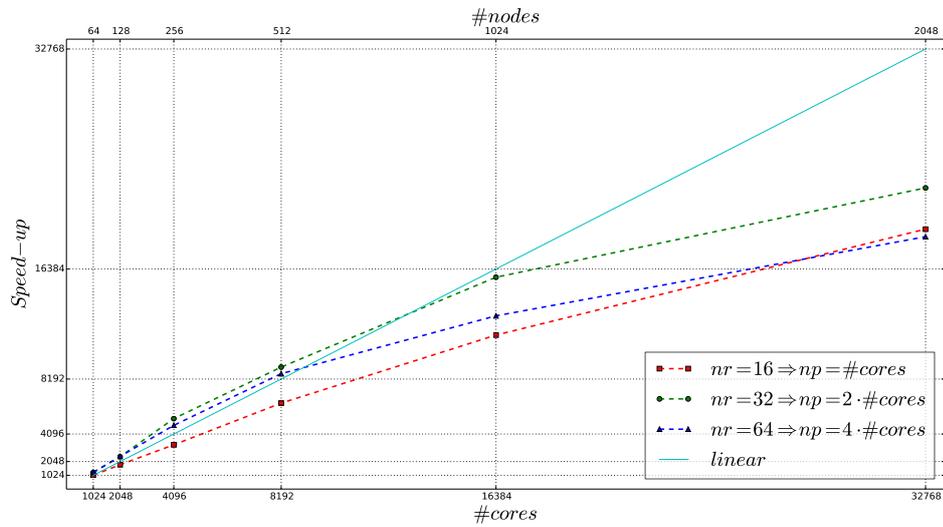


Figure 10.5: Scalability of the FDTD code: graph of the speed-up as a function of the MPI process number and core number. The unit of execution time is the one taken by a 1024 MPI processes simulation. Green and blue lines refer to 2 and 4 MPI processes per core respectively.

Usually an MPI process is executed on a distinct physical processor (core). This corresponds to the red line in Fig. 10.5 but, due to multithreading, each physical core can be overloaded with 2 or 4 distinct MPI processes, and these situation correspond to the green and blue lines in the Figure. To be noted is that this solution can, initially, outperform the linear speed-up increase. Another important aspect of the code parallelization is *load balancing*, which means that the number of floating point operations in the time unit performed by each MPI process should be equally distributed among them. Otherwise there is an inefficient use of the computational resources, which reflects in some processes staying idle, waiting for others to complete their task. To achieve load balancing the various spatial subdomains should be of equal size, but this could not suffice for a fine tuned balancing, because there can be an inherent diversity in the number of floating point operations due, for example, to Yee cells belonging to the outer absorbing layer of to disper-

sive materials, with respect to those belonging to conventional materials. Summarizing, one can say that in general, for a given model under study, there will be an optimal MPI process number and a related decomposition which assure the maximum efficiency. To pretend to increase at all costs that number means unavoidably a computational resources waste.

Chapter 11

Plasmonics applications

In the present Chapter the analytical results for the sphere, and the numerical FDTD simulations for different, non analytically solvable, shapes, are used to study and compare among them the scattering and absorption coefficients of small *metal* objects [49]. Different types of metal and different shapes and sizes are considered. From the electromagnetic point of view, the metal properties determining the response to light are included via a Drude-Sommerfeld + double Critical Points model [48], expressed quantitatively by equations (1.7), (1.10) and (1.13) of Chapter 1. This amounts to a classical description for the excitation of the so called Surface Plasmon (SP) modes supported by particles, i.e., metal free electron surface-charge oscillation resonances, occurring at well defined frequencies (vacuum wavelengths). A resonance condition occurs typically when the *real* part of the metal complex permittivity ε_c becomes negative, thus determining an almost zero condition of the denominator in the expression for the electric field or related quantities [18, 50].

The exploitation of SP resonances has been thought as a promising tool to enhance the efficiency of photovoltaic cells, through the presence of metal nanostructures in front or in the bulk of the cell themselves [51]. Typically, the light energy conversion in electricity should be enhanced by a far-field effect extending the optical path of the rays within the cell, according to the scheme of Fig. 11.1:

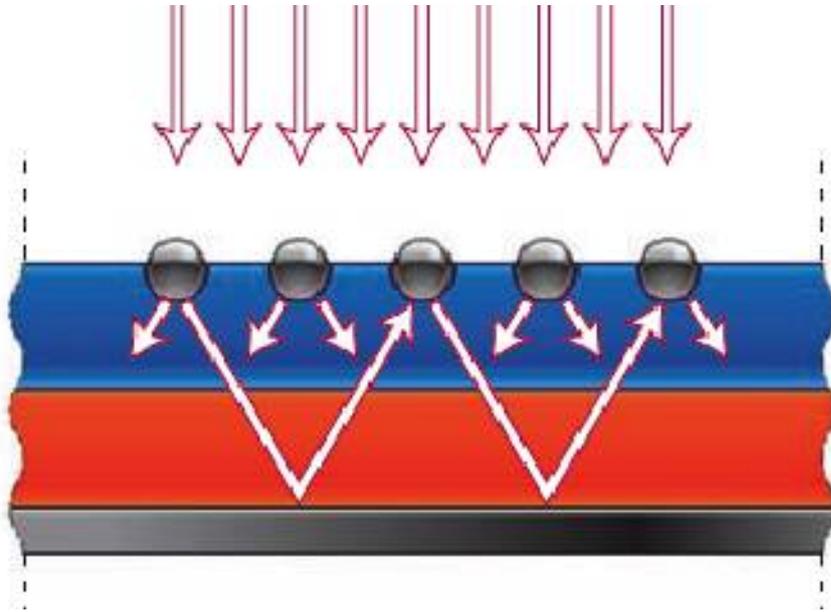


Figure 11.1: Optical path extension by scattering off nanoparticles in solar cells.

thus increasing the probability of charge carrier pairs production by each photon. Other proposed enhancement mechanisms through metal nanoparticles, but here not considered, are the near-field effect locally enhancing the energy conversion, or the creation of energy-rich charged carriers which are then transferred to the solar cell.

11.1 Scattering from spherical nanoparticles. Analytical case.

The present Section will directly take advantage of the analytical expressions (10.1) for the spectral coefficients of an isolated metallic sphere in a homogeneous non-absorptive medium. Considering more than ten terms in the series expansions gives accurate enough results, well beyond a simple dipole approximation which is valid, as pointed out in [52], if $2\pi r/\lambda \ll 1$ where r is the sphere radius. Metals considered in this Section are gold (Au), silver (Ag), aluminum (Al) and copper (Cu). The parameter values to be used in (1.10) and (1.13) for Au and Ag were already given previously. For Al they are, from [48]:

$$\begin{aligned}
\varepsilon_{r,\infty} &= 1.0000 \\
\omega_{\text{pl}} &= 2.0598 \times 10^{16} \text{ rad/sec} \\
\gamma &= 2.2876 \times 10^{14} \text{ rad/sec} \\
A_1 &= 5.2306 \\
\phi_1 &= -0.51202 \\
\Omega_1 &= 2.2694 \times 10^{15} \text{ rad/sec} \\
\Gamma_1 &= 3.2867 \times 10^{14} \text{ rad/sec} \\
A_2 &= 5.2704 \\
\phi_2 &= 0.42503 \\
\Omega_2 &= 2.4668 \times 10^{15} \text{ rad/sec} \\
\Gamma_2 &= 1.7731 \times 10^{15} \text{ rad/sec}
\end{aligned}$$

Here they are simply used to calculate the complex electric permittivity ε_c to be inserted in (10.1). Data for ε_c of Cu are got, for each frequency (vacuum wavelength) considered, from [53]. See also [54]. Furthermore, three sphere radiuses are considered: 50, 100, 150 nm. Finally, three *non-absorptive* embedding media are also taken into account: air, silica (SiO_2) and silicon nitride (Si_3N_4). Their purely real and constant permittivities are: 1, 2.34, 4.08 respectively. The different embedding media could reflect the presence of a dielectric passivating layer in the solar cell. The scattering coefficient $Q_{\text{scatt}}(\lambda)$ results in the range 200 ÷ 1100 nm are given in Fig. 11.2 below, as obtained from (10.1a) [49].

As stems out from the Figure, Q_{scatt} has an averaged value of $Q_{\text{scatt}}^* \approx 3 \div 4$ over the wavelength range, which implies a circular cross section area of radius $r\sqrt{Q_{\text{scatt}}^*}$. In an arrayed system of spherical particles they should thus stay at a typical distance $d > 2r\sqrt{Q_{\text{scatt}}^*}$ to avoid effective areas overlapping and consequent cooperative scattering effects. Such a constraint is compatible with a suggested surface coverage of 30% found in [55] (surface coverage means the percentage of the unit surface area filled with particles), because with the lower bound for the mutual distance given above, it results a required surface coverage of about $\pi/4 Q_{\text{scatt}}^* \approx 0.26$. This value also allows to limit the scattering and absorption analyses to a single isolated sphere. The absorption coefficient $Q_{\text{abs}}(\lambda)$ results in the range 200 ÷ 1100 nm are given in Fig. 11.3 below, as obtained by subtracting (10.1a) from (10.1b) [49].

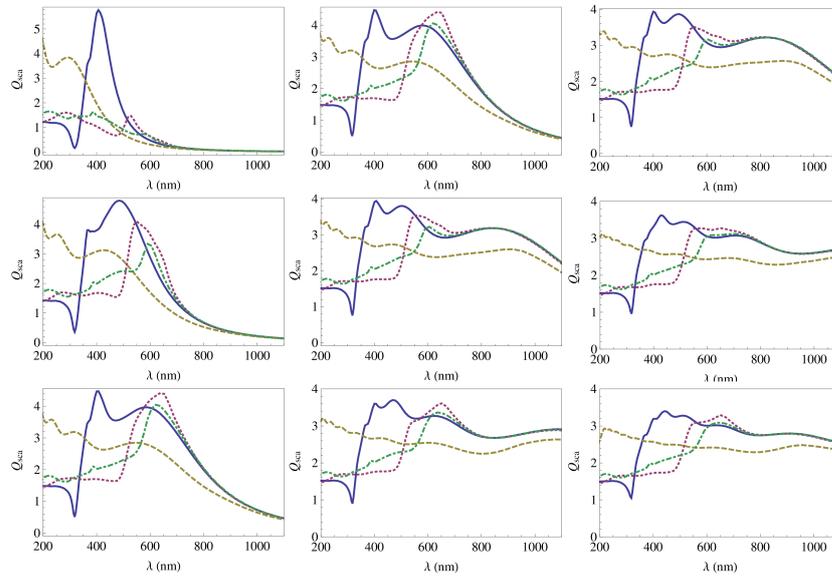


Figure 11.2: Scattering coefficient Q_{scatt} for an isolated sphere as a function of the wavelength λ . *Top row*: air embedded; *middle row*: SiO_2 embedded; *bottom row*: Si_3N_4 embedded. *Left column*: $r = 50$ nm; *central column*: 100 nm; *right column*: 150 nm. *Blue line*: Ag; *red dotted line*: Au; *yellow dashed line*: Al; *green dash-dotted line*: Cu.

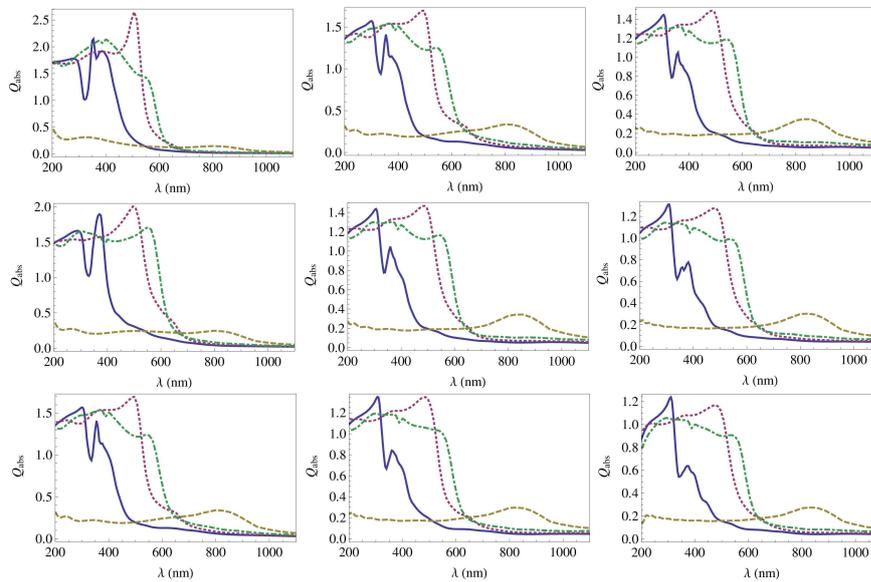


Figure 11.3: Absorption coefficient Q_{abs} for an isolated sphere as a function of the wavelength λ . *Top row*: air embedded; *middle row*: SiO_2 embedded; *bottom row*: Si_3N_4 embedded. *Left column*: $r = 50$ nm; *central column*: 100 nm; *right column*: 150 nm. *Blue line*: Ag; *red dotted line*: Au; *yellow dashed line*: Al; *green dash-dotted line*: Cu.

For the purposes at hand the scattering contributions should dominate over the absorption ones, thus a good parameter is the *albedo* α , defined as:

$$\alpha(\lambda) = 100 \frac{Q_{scatt}(\lambda)}{Q_{abs}(\lambda) + Q_{scatt}(\lambda)} = 100 \frac{Q_{scatt}(\lambda)}{Q_{ext}(\lambda)}.$$

It is shown in Fig. 11.4 below:

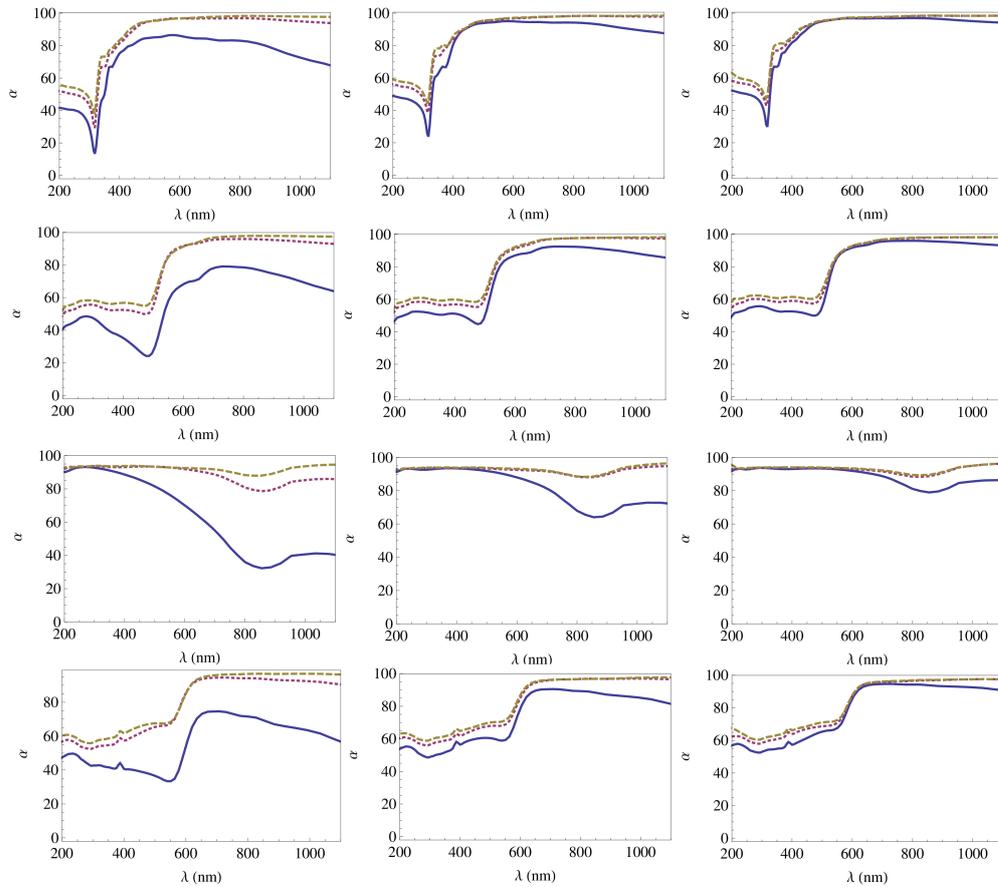


Figure 11.4: Albedo α for an isolated sphere as a function of the wavelength λ . *Top row*: Ag; *second row*: Au; *third row*: Al; *last row*: Cu. *Left column*: air embedded; *central column*: SiO₂ embedded; *right column*: Si₃N₄ embedded. *Blue line*: $r = 50$ nm; *red dotted line*: $r = 100$ nm; *yellow dashed line*: $r = 150$ nm.

To the spectral parameters just considered can be given a more comprehensive meaning if they are averaged over the whole wavelength range, weighting them with the spectral intensity of the solar light $I(\lambda)$, the standard AM1.5G spectrum intensity [56]:

$$\langle Q_s \rangle = \frac{\int I(\lambda) Q_s(\lambda) d\lambda}{\int I(\lambda) d\lambda}$$

where s stands for *ext*, *scatt*, *abs* and the integration in the present case has been numerically performed, with a trapezoidal formula, from 300 to 1100 nm. Such averaged values allow a quicker comparison when metal type, size and embedding medium are varied. Another figure of merit is the fraction of the total scattered power which is effectively scattered forward, because if the nanoparticles are placed on the front surface of a solar cell, one wants this radiation to be concentrated in the direction of the *substrate*. Thus it can be defined as:

$$f_{subs} = \frac{\int_0^{2\pi} \int_0^{\pi/2} \vec{P}_{scatt} \cdot \hat{u}_r \sin(\theta) d\theta d\phi}{\int_0^{2\pi} \int_0^{\pi} \vec{P}_{scatt} \cdot \hat{u}_r \sin(\theta) d\theta d\phi}$$

where \vec{P}_{scatt} is the real part of the complex Poynting vector for the scattered field off the sphere, which can be calculated using the series expansions (1.14), with the expansion coefficients (1.15a) and (1.15b), which are also used in (10.1). \hat{u}_r is the radial unit vector and the spherical coordinate system is chosen in such a way that the half solid angle considered in the numerator of the expression above is the one pointing in the same direction of the impinging light beam. Also for f_{subs} could be considered the averaged value $\langle f_{subs} \rangle$ with the weight function $I(\lambda)$ as above.

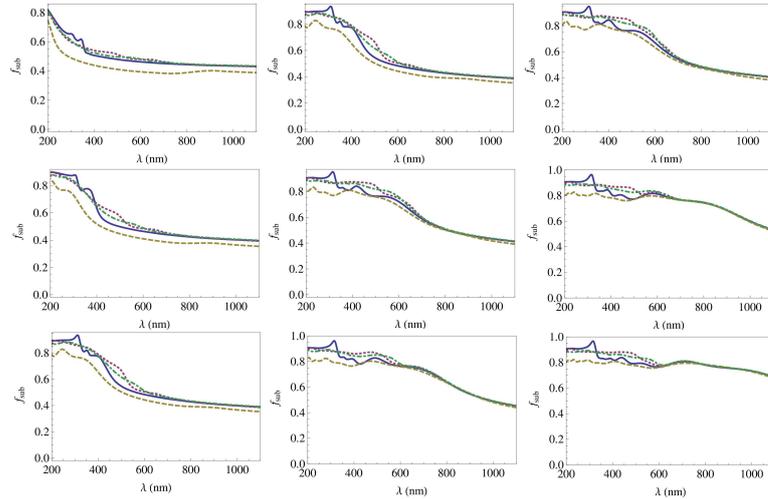


Figure 11.5: Fraction of forward scattered off sphere light as a function of the wavelength λ . *Top row*: air embedded; *second row*: SiO_2 embedded; *third row*: Si_3N_4 embedded. *Left column*: $r = 50$ nm; *central column*: $r = 100$ nm; *right column*: $r = 150$ nm. *Blue line*: Ag; *red dotted line*: Au; *yellow dashed line*: Al; *green dash-dotted line*: Cu.

11.1. SCATTERING FROM SPHERICAL NANOPARTICLES. ANALYTICAL CASE.117

• As can be seen from Fig. 11.2, the scattering in air is better for $r = 150$ nm, but with SiO_2 or Si_3N_4 embeddings, a radius value toward $r = 100$ nm performs better. Larger spheres exhibit more than one resonance peak, meaning multipolar excitation, with high Q_{scatt} values spread on a larger part of the spectrum. From Fig. 11.2 one can see that the presence of an external dielectric medium increases the scattering at long wavelengths, but by reducing the maximum value and widening the profile as a function of λ . This is due to the dielectric polarization which partially screens the metal surface charges thus weakening the restoring force among them. Ag is by far the best scatterer, followed by Al, Au and Cu. The considerations above are condensed in Tab. 11.1 for the averaged $\langle Q_{scatt} \rangle$ values over the range $300 \div 1100$ nm.

$r = 50$ nm				
	Ag	Au	Al	Cu
air	0.98	0.49	0.60	0.41
SiO_2	2.19	1.74	1.70	1.40
Si_3N_4	2.82	2.45	2.28	2.12

$r = 100$ nm				
	Ag	Au	Al	Cu
air	2.80	2.44	2.27	2.10
SiO_2	3.20	2.85	2.55	2.72
Si_3N_4	3.09	2.76	2.48	2.74

$r = 150$ nm				
	Ag	Au	Al	Cu
air	3.21	2.84	2.56	2.71
SiO_2	3.03	2.72	2.47	2.69
Si_3N_4	2.95	2.67	2.42	2.63

Table 11.1: Averaged scattering coefficients $\langle Q_{scatt} \rangle$ for metal spheres embedded in different media.

• For what concerns absorption, Fig. 11.3 points out that it is reduced by an increasing of the size and/or the relative permittivity of the external dielectric, with Cu and Al less affected by this rule. In particular, Ag and Al are very poor absorbers, with a high albedo. It is maximum, in general, for $r = 150$ nm and/or for Si_3N_4 embedding, as results from Fig. 11.4. Au and

Cu absorb more light, especially at the peak of the solar spectrum between 400 and 500 nm, as can be seen from Fig. 11.3. This circumstance is a drawback for photovoltaic applications.

- Regarding the angular distribution of the scattered radiation, Fig. 11.5 shows that shorter wavelengths are forward scattered more than the longer ones. Also, as is well known, spheres with larger size tend to scatter more in the forward direction. From Fig. 11.5 one can see that the same behavior is noted when the relative electric permittivity of the embedding medium is increased, due to the excitation of higher order multipoles, which reemitt in the forward direction more than a dipole. The graphs of Fig. 11.5 are almost the same for different materials, although Al is less efficient than the other metals.

11.2 Scattering from nonspherical nanoparticles. FDTD case.

In the present Section, Al isolated nanoparticles embedded in air and of different sizes and shapes are considered for an FDTD numerical analysis akin to the one of the previous Section. In particular, ellipsoidal particles with a total height of $h = 100$ (minor axis) and $h = 300$ nm (major axis), both with a circular projection of radius $r = 100$ nm in a plane normal to the incident radiation are studied, but also an hemispherical particle with $r = h = 100$ nm and cylindrical particles of height $h = 100, 200, 300$ nm with a circular base of radius $r = 100$ nm are simulated. Fig. 11.6 shows intensity color maps of the electric field as obtained by means of the FDTD calculations, done with with a lattice of $120 \times N_y \times 120$ cubic Yee cells of edge size $\delta_x = \delta_y = \delta_z = 2$ nm and with a varying number of cells along the y -axis, the beam propagation direction: $N_y = 120, 160, 216$ to accomodate the different particles.

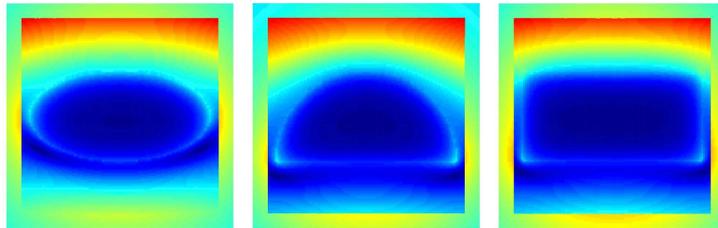


Figure 11.6: Color maps of the electric field intensity inside and surrounding Al ellipsoid, hemisphere and cylinder, all of height $h = 100$ nm. Radiation is propagating from the top to the bottom sides of the figures. The field is linearly polarized in the direction normal to the page. Colors are in logarithmic scale of intensity, normalized to a reference value of 1 V/m. The maps refer to a vacuum wavelength of 500 nm.

11.2. SCATTERING FROM NONSPHERICAL NANOPARTICLES. FDTD CASE.119

Fig. 11.7 below compares Q_{scatt} , Q_{abs} and the albedo α for the different shapes and sizes to the “reference” sphere with $r = 100$ nm, at the various wavelengths:

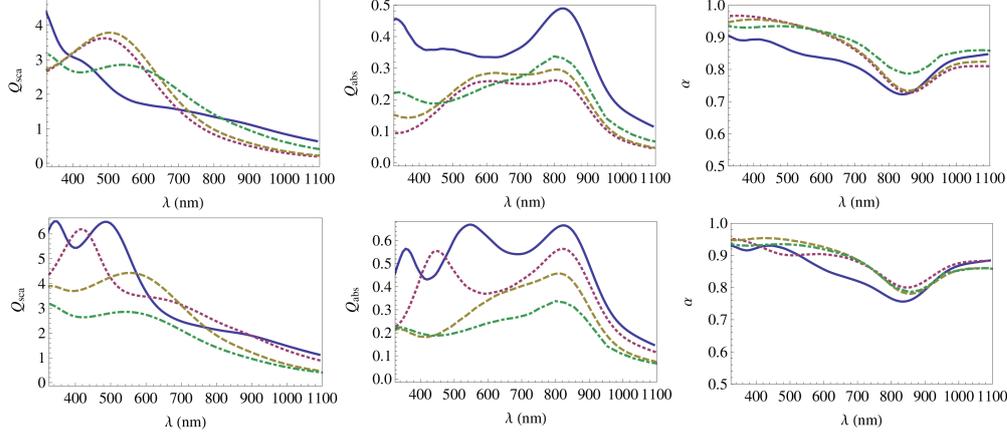


Figure 11.7: Scattering coefficients (*left*), absorption coefficients (*middle*) and albedo (*right*) as functions of the wavelength for an aluminum nanoparticle. *First row*, blue line: ellipsoid with $h = 300$ nm; red dotted line: ellipsoid with $h = 100$ nm; yellow dashed line: hemisphere with $r = 100$ nm; green dash-dotted line: sphere with $r = 100$ nm. *Second row*, blue line: cylinder with $h = 300$ nm; red dotted line: cylinder with $h = 100$ nm; green dash-dotted line: sphere with $r = 100$ nm. All particles embedded in air.

while Tab. 11.2 below shows the averaged values of the scattering and absorption coefficients over the $300 \div 1100$ nm range.

	h (nm)	$\langle Q_{scatt} \rangle$	$\langle Q_{abs} \rangle$
Cylinders	100	2.98	0.30
	200	3.29	0.41
	300	3.57	0.53
Ellipsoids	100	2.00	0.20
	300	1.75	0.35
Hemisphere	100	1.81	0.22
Sphere	200	2.27	0.24

Table 11.2: Averaged scattering and absorption coefficients $\langle Q_{scatt} \rangle$ and $\langle Q_{abs} \rangle$ for metal nanoparticles of different shapes and sizes embedded in air.

Spheroidal shapes (including the hemisphere) have similar behaviors. Abandoning the purely spherical shape exhibits a shift of the scattering peaks, while increasing the height increases the absorption. On the other hand cylinders, independently from their height, have increased scattering and absorption with respect to the sphere, both at each single wavelength λ and

as integrated values, even if the albedo resembles closely that of the sphere. Cylinders thus seem to be a better choice as scatterer.

A numerical implementation of the Kirchhoff integral formula (2.16) is also used to calculate the integral in the definition of $f_{subs}(\lambda)$, using a radiation amplitude distribution sampled at steps of 1° . The resulting graphs are shown in Fig. 11.8 below:

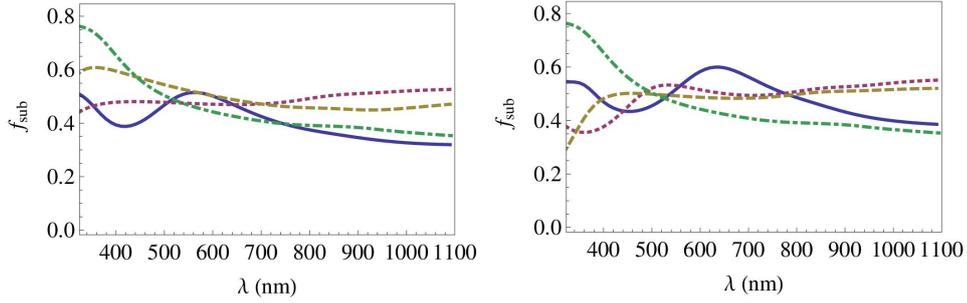


Figure 11.8: Fraction of light forward scattered off nanoparticles as a function of the wavelength λ . *Left blue line:* ellipsoid with $h = 300$ nm; *red dotted line:* ellipsoid with $h = 100$ nm; *yellow dashed line:* hemisphere with $r = 100$ nm; *green dash-dotted line:* sphere with $r = 100$ nm. *Right blue line:* cylinder with $h = 300$ nm; *red dotted line:* cylinder with $h = 200$ nm; *yellow dashed line:* cylinder with $h = 100$ nm; *green dash-dotted line:* sphere with $r = 100$ nm. All particles in air.

while the integrated or averaged values over the $300 \div 1100$ nm range are given in Tab. 11.3:

	h (nm)	$\langle f_{subs} \rangle$
Cylinders	100	0.49
	200	0.50
	300	0.49
Ellipsoids	100	0.49
	300	0.42
Hemisphere	100	0.50
Sphere	200	0.45

Table 11.3: Averaged fraction of light forward scattered off nanoparticles.

As can be seen from Fig. 11.8, cylinders and spheroids are less performing than the sphere at shorter wavelengths, but they tend to perform better at the longer ones. If the averaged values are considered, Tab. 11.3 indicates that the sphere is outperformed by all other shapes, excluding the ellipsoid with $h = 300$ nm.

11.3 Arrayed particles on a silicon substrate.

In this Section the numerical FDTD analysis is used for the optical characterization of an array system of hemispherical silver (Ag) particles embedded in a SiO_2 (silica) layer, and then deposited on a silicon (Si) substrate. This study is made in view of the possible enhancement in solar cells efficiency, by exploiting surface plasmonic effects in the nanoparticles to increase the overall fraction of light reaching the active part of the photovoltaic device. The presence of a silica coating with a typical thickness of 100 nm on top of crystalline silicon wafers is common for anti-reflection purposes (ARC), but also to passivate the Si surface and reduce electron-hole recombination. In the present Section an hybrid ARC is proposed because, as has been seen in previous Sections 11.1 and 11.2, an higher dielectric constant environment enhances the forward scattering off the nanoparticles and simultaneously they are mechanically and chemically protected and, furthermore, they are not in contact with the silicon surface, thus preventing an increase in surface carrier recombination. The numerical FDTD simulations are here aimed to compare in a systematic way the transmission properties of the proposed hybrid ARC with that of the standard silica coating and with configurations where the metal Ag hemispherical nanoparticles are placed on bare silicon or above the dielectric layer. Various parameters are considered in the simulations: the size of the hemispherical particles (radius), the surface coverage and the thickness of the dielectric layer above the nanoparticles. Below them there is always a silica bed as a separation from the Si surface. The silica coating is also assumed conformal, as suggested by Fig. 11.9:

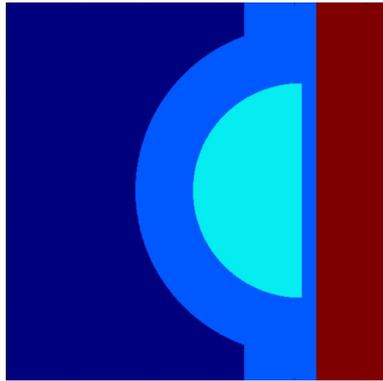


Figure 11.9: Schematic of the plasmonic Ag hemisphere (light blue) embedded in the 100 nm thick SiO_2 layer (blue), on top of a crystalline Si wafer (brown). On the left of embedding is vacuum (dark blue).

and presenting bumps in correspondence of the hemispherical particles. The profile in Fig. 11.9 is believed to be more realistic than a simpler flat one or

with other curvatures. The arraying of the Ag particles is FDTD simulated by means of periodic boundary conditions (PBC) as previously explained in Chapter 6 (Section 6.2) and the transverse dimensions along the coordinate axes to which the PCB are applied are chosen as a function of the surface coverage. Along the impinging light beam propagation direction, CPML boundary conditions are applied (Section 6.1), one side being in free space, the other just at the bottom of Si substrate which, as a consequence of the absorbing layer behavior, can be considered very large (theoretically infinite). For the present Section are also available experimental measurements of reflectivity, obtained by means of a photo spectrometer coupled with an integrating sphere (Varian Cary 5000). The measurements were made on produced samples with hybrid coating. The fabrication process is schematically described in Fig. 11.10 below:

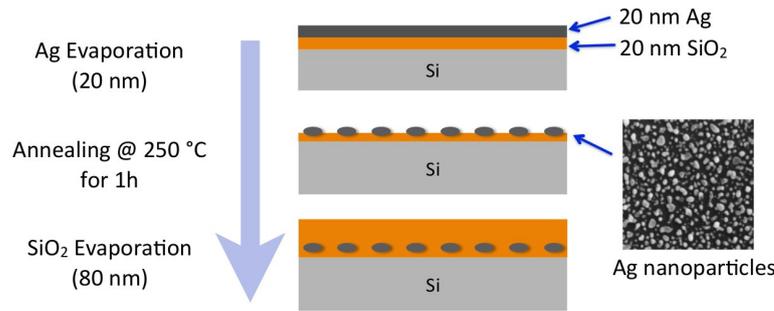


Figure 11.10: Schematic representation of the fabrication procedure of the photovoltaic cell with the embedded Ag nanoparticles in the SiO₂. In the inset also a Scanning Electron Micrograph (SEM) picture of the produced test samples.

The ensemble of Ag nanoparticles is obtained through the intermediate deposition of an Ag precursor layer of thickness 15 nm or 20 nm and subsequent annealing. As can be appreciated from the SEM micrograph inset in Fig. 11.10, the particle distribution is irregular, mainly for what concerns the particle shapes and sizes. The simulations on the other hand assume mono-sized particles regularly arrayed and with a well defined shape, and these differences should be taken into account when the simulated and measured data results are compared. Both, the fabrication of the photovoltaic cell and the reflectivity measurement were performed at MNF, the Micro-Nano characterization and fabrication Facility of Fondazione Bruno Kessler (FBK), Centre for Materials and Microsystems (CMM).

The optical properties of crystalline silicon, i.e., its angular frequency dependent relative complex permittivity, are described according to [45]:

$$\varepsilon_r(\omega) = \varepsilon_{r,\infty} + \sum_{\ell=1}^2 \frac{\Delta_{\ell} (\Omega_{\ell}^2 - i\Gamma_{\ell}\omega)}{\Omega_{\ell}^2 - 2i\gamma_{\ell}\omega - \omega^2} \quad (11.1)$$

where i is the imaginary unit, which is amenable to the expression (8.1) given in Chapter 8, and which is included in the FDTD algorithm as there described. The choice (11.1), with the following values for its parameters [45]:

$$\begin{aligned} \varepsilon_{r,\infty} &= 1 \\ \Delta_1 &= 8.93 \\ \Delta_2 &= 1.855 \\ \Omega_1 &= 3.42 \times (2\pi c_o \times 10^6) \text{ rad/sec} \\ \Omega_2 &= 2.72 \times (2\pi c_o \times 10^6) \text{ rad/sec} \\ \Gamma_1 &= 0.087 \times (2\pi c_o \times 10^6) \text{ rad/sec} \\ \Gamma_2 &= 2.678 \times (2\pi c_o \times 10^6) \text{ rad/sec} \\ \gamma_1 &= 0.425 \times (2\pi c_o \times 10^6) \text{ rad/sec} \\ \gamma_2 &= 0.123 \times (2\pi c_o \times 10^6) \text{ rad/sec} \end{aligned}$$

where c_o is the speed of light in vacuum, is justified by an improved fitting with the experimental data for both the visible and near ultraviolet ranges. It can be seen from the following Fig. 11.11:

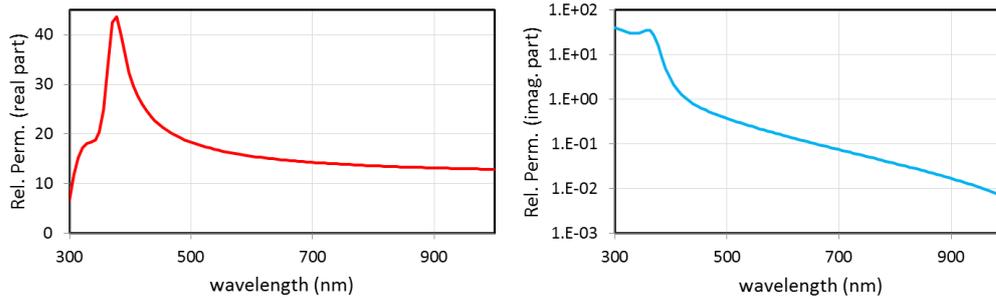


Figure 11.11: Real (left) and imaginary (right) parts of Si relative complex permittivity according to (11.1) with the above parameters [45], on the 300 ÷ 1000 nm wavelength range. In the right graph the vertical axis has a logarithmic scale.

that below 400 nm Si exhibits a quick increase of its dielectric constant and some non-negligible absorption.

In Fig. 11.12 it is shown a comparison among experimental, analytical and

FDTD simulated reflectance of a standard pure silica ARC layer (100 nm thick) on top a silicon wafer:

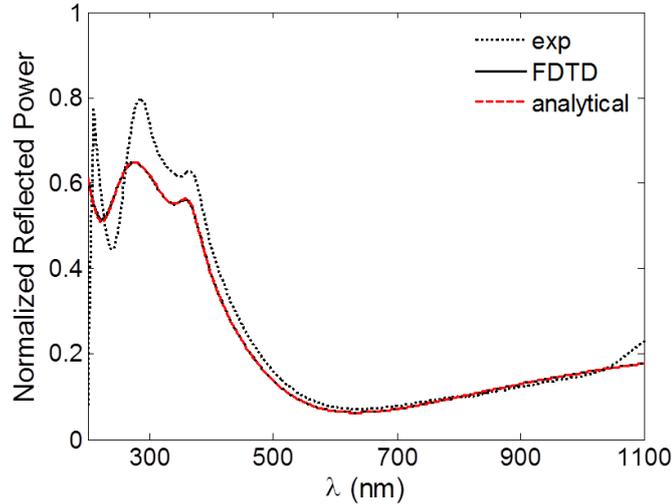


Figure 11.12: Comparison among experimental, analytical and FDTD numerical spectral reflectances for a 100 nm thick SiO_2 homogeneous antireflecting coating (ARC) layer on top of a $300 \mu\text{m}$ thick Si wafer.

in which it is apparent a fairly good agreement. The analytical one is obtained by means of the calculations explained in the Chapter on planarly multilayered media (Chapter 4, Section 4.1). In the analytical and in the FDTD calculations, the Si relative electric permittivity of Eq. (11.1) has been used. In particular, in the FDTD ones, the formalism for dispersive materials as described in Chapter 8, Section 8.2, is used. The small deviation in the experimental curve at about 1050 nm is ascribable to additional back-reflection due to the finite thickness of the Si substrate. It should be also pointed out that FDTD results for this configuration do not require a full 3D simulation, but only an 1D one.

Fig. 11.13 shows the experimental reflectance spectra for various hybrid layers as compared with the standard ARC made with silica alone and 100 nm thick used as reference. The first number before the slash sign in the legend indicates the thickness of the precursor Ag layer before its annealing, and the subsequent numbers (after the slash sign) are the thicknesses of the silica layers deposited before and after the Ag annealing. The effect of the nanoparticles is to decrease the reflectivity in a wavelength region below 450 nm and to increase it at higher wavelengths, the effect of the 20-80 configuration being slightly more pronounced than that of the 10-90. The plasmonic reflectance is peaked around 350 nm.

Fig. 11.14 compares the spectral reflectance experimental results for a precur-

sor layer 20 nm thick, with the FDTD simulated ones for various monosized particles radii and a coverage $\eta = 25\%$. The simulated spectra for monosized particles are significantly more structured than the experimental ones, and it appears like if these latter could be understood as superpositions of size-distributed Ag nanoparticles (as is indeed the actual situation).

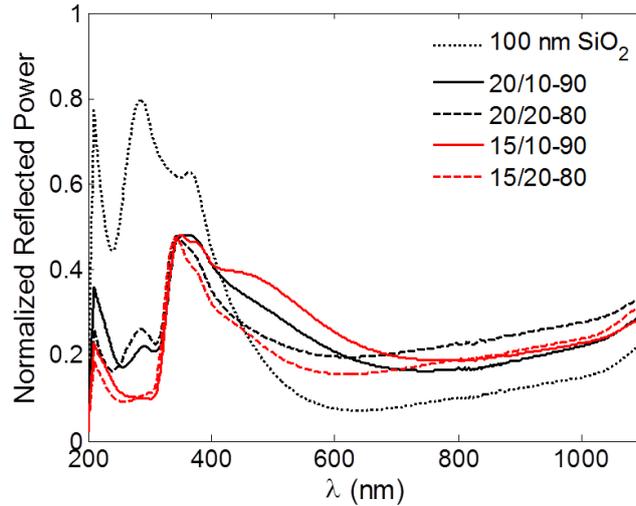


Figure 11.13: Experimentally measured spectral reflectances from a 100 nm thick SiO_2 layer on top of a $300 \mu\text{m}$ thick Si wafer (dotted curve), along with several configurations of plasmonic Ag particle arrays embedded in the silica layer as indicated in the legend (see also text).

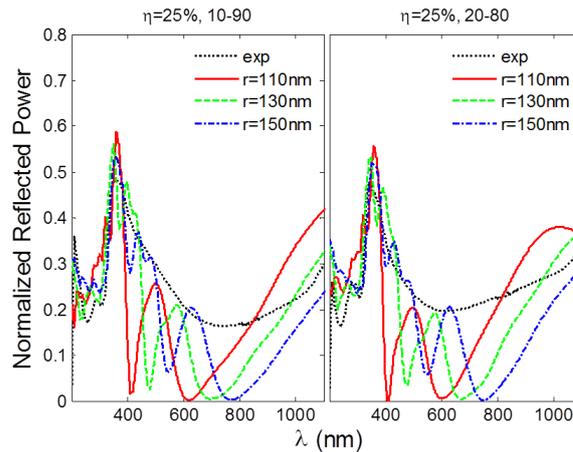


Figure 11.14: Measured and simulated spectral reflectances from Ag plasmonic particle arrays embedded in a SiO_2 layer in the 10-90 (left panel) and 20-80 configurations. Experimentally measured curves (dotted) refer to samples prepared via deposition of a 20 nm thick Ag layer. Simulated curves are calculated for a coverage $\eta = 25\%$ and for an hemisphere radius $r = 110$ nm (continuous), $r = 130$ nm (dashed) and 150 nm (dash-dotted).

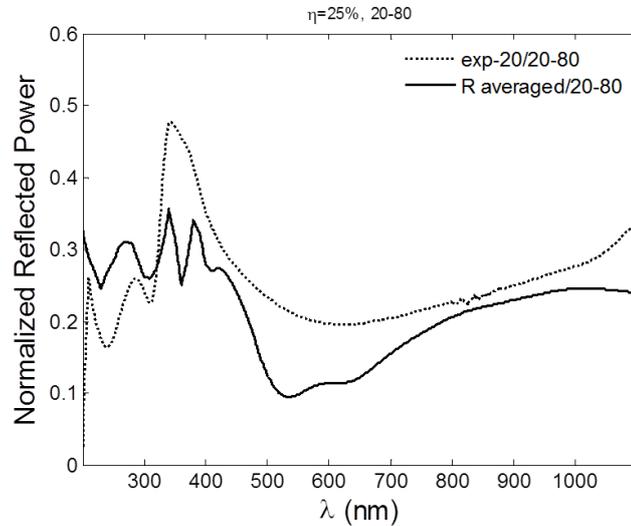


Figure 11.15: Averaged FDTD calculated reflectance for an array of Ag embedded particles as compared with the measure one. Configuration is 20-80 and with an assumed coverage $\eta = 25\%$.

Fig. 11.15 shows an averaged calculated reflectance spectrum obtained from the FDTD reflectance values, after an averaging procedure. The weights of the averaging were extracted from a discrete population distribution of the Ag particle sizes. These sizes were evaluated at FBK-MNF by means of a dedicated software, which fitted the particle images analyzing SEM pictures from the fabrication process. The Figure is aimed to demonstrate that a distribution of particles of different sizes gives an undifferentiated spectrum contrasting with the typical ones of monosized particles, which are characterized by typical resonance peaks.

The previous comparisons with experimental data involved reflectance spectra because the measurements could only be obtained by detecting the reflected power. In what follows instead, a series of transmittance spectra comparisons of FDTD simulation results with the standard silica ARC, assumed as reference response curve, are reported. The transmitted power in the FDTD simulations is calculated by integrating the Poynting vector on a surface just beneath the silica/silicon interface, a few FDTD unit cells apart from it. The spectral data are corrected by the calculated absorption in the thin slab between the silica/silicon interface and the integration plane. In fact, when the behavior of the standard ARC without particles is calculated, such absorption values are recorded. These are then subtracted from the absorbed ones and added to the transmitted ones, when the behavior of the Ag nanoparticles is evaluated.

Fig. 11.16 shows the transmittance for monosized hemispherical particles of radius $r = 150$ nm and for a 10-90 configuration, parameterized with a vary-

ing coverage η . It results that, starting from $\eta = 20\%$, two peaks appears, the higher at 800 nm and which does not shift with η , the other moving from 600 nm to 500 nm as the coverage increases. A dip between the peaks becomes more and more pronounced as η increases. An interesting window opens, from 730 nm to 960 nm, in which the transmittance is higher than that of the standard ARC. Its positioning on the wavelength range is however not so well suited with respect to the solar intensity distribution spectrum.

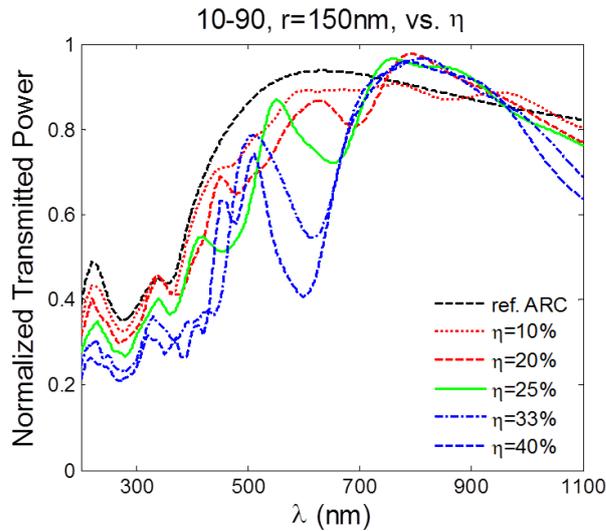


Figure 11.16: Simulated spectral transmittances for an array of hemispherical monosized Ag particles of radius $r = 150$ nm, embedded in a SiO_2 layer in the 10-90 configuration, at different coverages. Also shown the transmittance of the ARC alone (black dashed).

The next graphs in Fig. 11.17, in which the parameterization involves the particle size and its distance from the silica/silicon interface (with $\eta = 25\%$), confirm that for a radius $r = 130$ nm or greater, and with the smaller separation from the interface of 10 nm, the plasmonic hybrid ARC is able to outperform the standard ARC in the restricted range $730 \div 960$ nm, even if its broadband performance remains still worse than that of a standard ARC. In Fig. 11.18 spectral FDTD calculated transmittances are shown for a broader set of particle sizes and for both 10-90 and 20-80 configurations, with coverage $\eta = 25\%$, as compared with the standard ARC. Superimposed to the graphs there is also the solar AM1.5G spectrum.

In Fig. 11.19 spectral simulated transmittances for an increasing thickness of the covering silica layer above the hemispherical Ag particles, which are held at 10 nm from the silica/silicon interface. With 10-0 is indicated a configuration in which there is only the Ag particles supporting silica layer 10 nm thick, but nothing covering them. In this configuration the spectrum

exhibits only a single peak, at around 580 nm for $r = 150$ nm and at around 510 nm for $r = 130$ nm, while particles embedding lead to the appearance of two maxima.

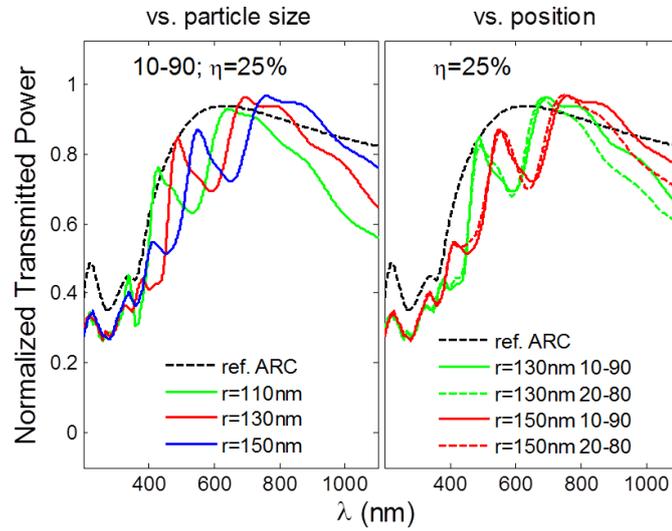


Figure 11.17: Simulated spectral transmittances for an array of hemispherical monosized Ag particles at a coverage of $\eta = 25\%$, in the 10-90 configuration and with different radii (left panel), and at different positions inside the SiO_2 layer (right panel). On both panels the transmittance of the ARC alone is shown (black dashed).

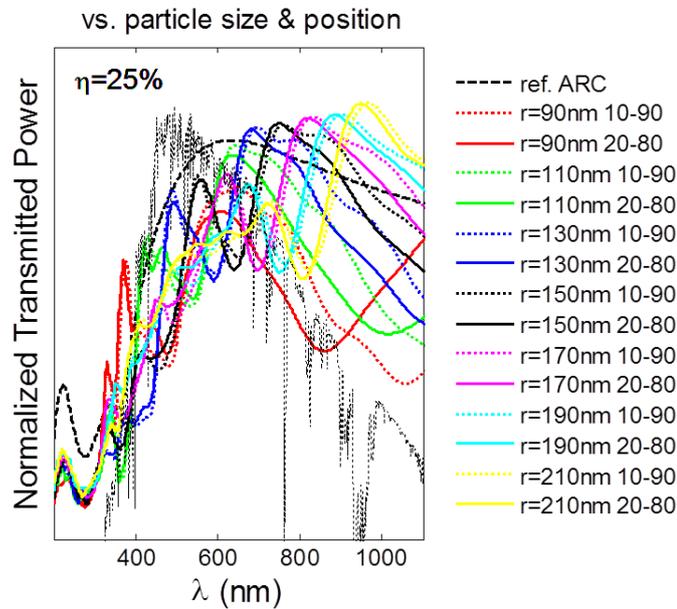


Figure 11.18: Simulated spectral transmittances for an array of hemispherical monosized Ag particles at a coverage of $\eta = 25\%$, in both the 10-90 and 20-80 configurations, with different radii. The transmittance of the standard ARC is shown (black dashed) along with the AM1.5G solar spectrum.

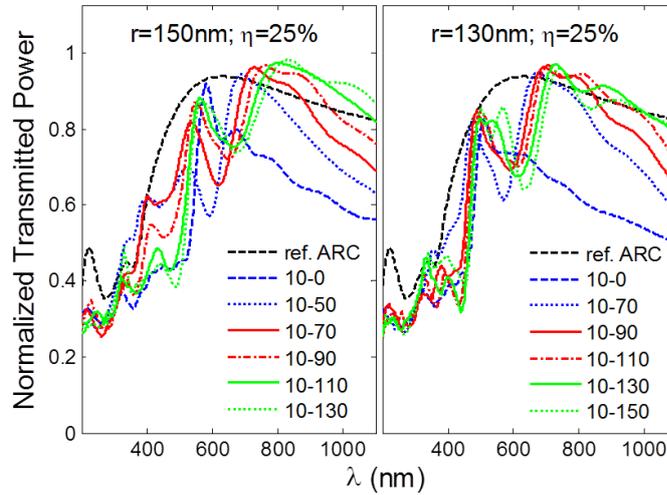


Figure 11.19: Simulated spectral transmittances for an array of hemispherical monosized Ag particles at a coverage of $\eta = 25\%$, for radii $r = 150$ nm (left panel) and $r = 130$ nm (right panel), at different overall thicknesses of the embedding SiO_2 layer, all in a 10- X configuration. Also shown the spectral transmittances of the ARC alone (black dashed).

In Fig. 11.20 spectral simulated transmittance with and without embedding are directly compared. In the right panel, for the $r = 130$ nm case it is also shown the transmittance spectrum for the 100-0 configuration, i.e., a single supporting layer 100 nm thick with no covering.

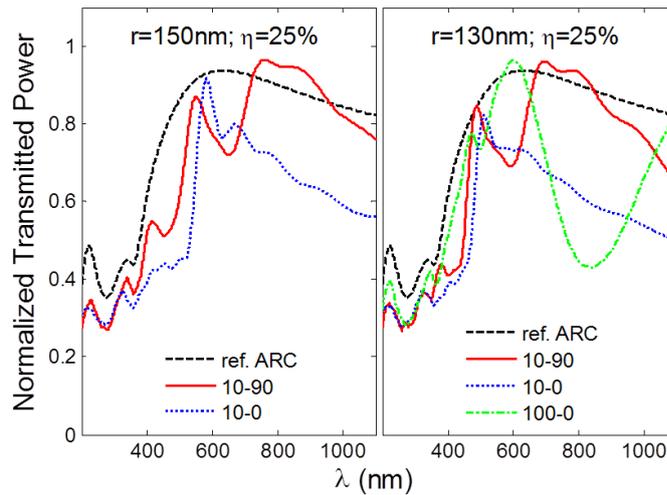


Figure 11.20: Simulated spectral transmittances for an array of hemispherical monosized Ag particles at a coverage of $\eta = 25\%$, for radii $r = 150$ nm (left panel) and $r = 130$ nm (right panel), where the particles are simply supported by, or embedded within, the SiO_2 layer. Also shown the spectral transmittances of the ARC alone (black dashed).

In Fig. 11.21 are reported averaged weighted transmittances on the wavelength range $200 \div 1100$ nm, where as weighting function is used the AM1.5G solar spectrum intensity, as functions of the coverage and the over-oxide thickness (under-oxide is always 10 nm). In Fig. 11.22 as functions of the hemisphere radius ($\eta = 25\%$).

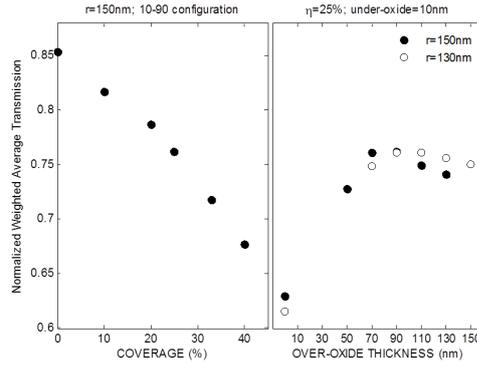


Figure 11.21: Simulated averaged transmittances weighted with the AM1.5G solar spectrum, as functions of the surface coverage (left panel) and of the over-oxide thickness grown on top of the Ag hemispherical particles (right panel).

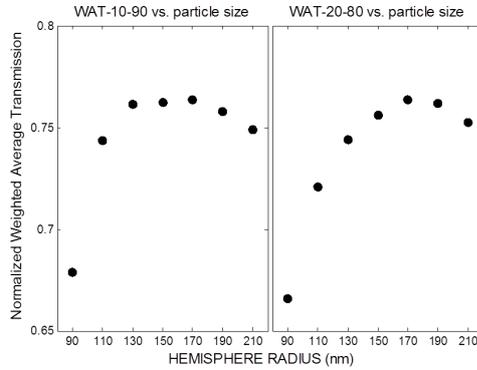


Figure 11.22: Simulated averaged transmittances weighted with the AM1.5G solar spectrum, as functions of the hemisphere radius.

It can be seen that the weighted transmittance decreases monotonically with increasing coverage, and with two characteristic rates (two slopes are appreciable in the left panel). This could be the clue that with a reduced free space among the Ag particles, cooperative effects between them can intervene. On the other hand, as a function of the oxide thickness grown on top of the Ag particles, transmittance goes through a maximum.

Finally, in Fig. 11.23 are shown some color maps of the electric field calculated for the array of Ag embedded nanoparticles with $r = 150$ nm (20-80

configuration and a coverage $\eta = 25\%$) at various frequencies — expressed as vacuum wavelengths — of the impinging monochromatic light. The view is a lateral one.

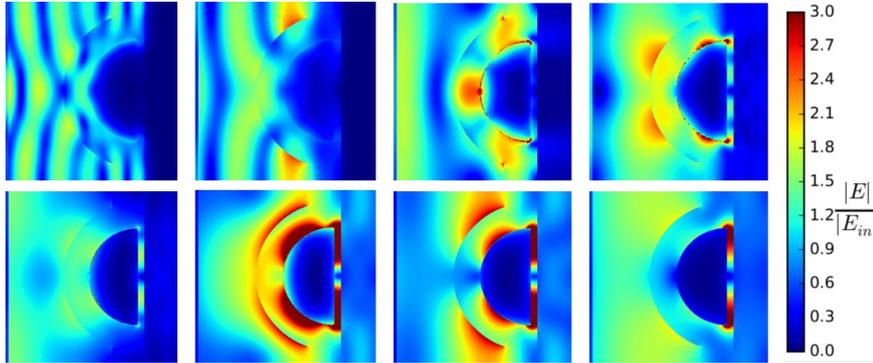


Figure 11.23: A sequence of electric field module maps corresponding to the vacuum wavelengths of 200 nm, 300 nm, 400 nm, 500 nm, 600 nm, 700 nm, 800 nm and 1000 nm (from top to bottom and from left to right) of the incident monochromatic light. Lateral view. Field values normalized to the incident ones.

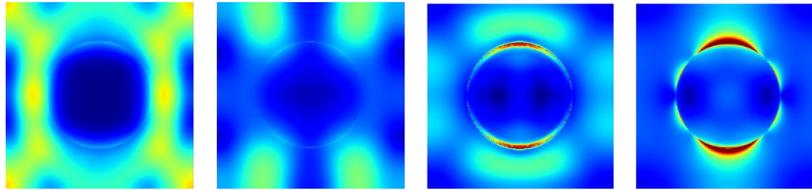


Figure 11.24: A sequence of electric field module maps corresponding to the vacuum wavelengths of 200 nm, 300 nm, 400 nm, 700 nm (from left to right) of the incident monochromatic light. Front view. The chromatic scale is the same as above. Field values normalized to the incident ones.

In Fig. 11.24 the view is a frontal one, corresponding to an observer watching in the direction of the impinging light beam. In both the electric field strength is normalized to that of the incident field.

As a concluding remark, in any case the broadband performances of the standard silica ARC remain unbeaten. The hybrid plasmonic ARC does seem to be useful in specific and selective applications, for example in sensoristic ones, within a restricted window of the overall $200 \div 1100$ nm wavelength range.

Chapter 12

Photonics application to an opal crystal

One of the main applications of the in-house FDTD code developed for the present work is in studying an opal photonic crystal interacting with a beam of light in the wavelength range of $400 \div 750$ nm ($1 \text{ nm} = 10^{-9} \text{ m}$). Photonic crystals [16] are the optical analogue of atomic or molecular crystals studied quantum mechanically in solid state physics, but with these microscopic constituents replaced by macroscopic media with differing electric permittivities. Their characteristic dimensions have to be comparable with the wavelength of the light. In common, both types of crystal have an ordered, periodic and symmetrical structure. In the case of the photonic crystals, instead of the force potential acting on the electrons, it is the electric permittivity function $\varepsilon(\vec{r})$ which exhibits translational and rotatational symmetries according to the underlying ordered structure. Because of the many similarities between the governing equations, the resulting Schrödinger wavefunction for electrons and the electromagnetic field inside the crystal give raise to similar physical phenomenologies. The most typical one is the existence of *photonic band gaps* preventing light from propagating with certain frequencies along certain directions. In particular, a *complete photonic band gap* stops the propagation of electromagnetic waves with any polarization in any direction. Such characteristics of photonic crystals make them useful for optical control and manipulation in analogy on how metallic waveguides and cavities control microwave propagation or confinement. To accomplish such tasks in the best way the dissipation of a photonic crystal has to be as low as possible, i.e. ε has to be, almost ideally, purely real (no imaginary part).

12.1 Eigenvalue problem and the band structure

From the theoretical point of view, the interaction of the light electromagnetic field with a photonic crystal amounts to the solution of the following *eigenvalue* problem:

$$\vec{\nabla} \times \left[\frac{1}{\varepsilon(\vec{r})} \vec{\nabla} \times \vec{H}(\vec{r}) \right] = \left(\frac{\omega}{c_o} \right)^2 \vec{H}(\vec{r}) \quad (12.1)$$

($c_o = \frac{1}{\sqrt{\mu_o \epsilon_o}}$, the speed of light in vacuum) with the further constraint:

$$\vec{\nabla} \cdot \vec{H}(\vec{r}) = 0.$$

Equation (12.1) is obtained by eliminating \vec{E} from the two curl equations (1.1c) and (1.1d) after a Fourier transform from the time domain to the frequency domain. Here as usual, but with an abuse of notation, the same letters \vec{E} and \vec{H} are used for a field variable as a function of time t or as a function of (angular) frequency ω . The time harmonic factor is $e^{-i\omega t}$. In (12.1) the *eigenvalue* is $(\omega/c_o)^2$, while the left member is a vector, differential, second order operator $\hat{\Theta}$ applied to the corresponding *eigensolution* $\vec{H}(\vec{r})$. In $\hat{\Theta}$, ε cannot be moved further to the left due to its spatial dependence which reflects, first of all, the periodic crystal structure:

$$\varepsilon(\vec{r} + \ell \vec{a}_1 + m \vec{a}_2 + n \vec{a}_3) = \varepsilon(\vec{r})$$

where $\ell, m, n \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ and \vec{a}_1, \vec{a}_2 and \vec{a}_3 are the *primitive lattice vectors* which span the elementary unit cell, whose tiling in \mathbb{R}^3 defines such underlying periodic structure. $\hat{\Theta}$ results to be an Hermitian operator with respect to a standard inner product for the solution space (but with a scalar dot product replacing the usual simple product between functions, which are now vector functions). If from (1.1c) and (1.1d) an equation for \vec{E} were deduced, instead of (12.1), the resulting eigenproblem would be a *generalized* one (with operators on both sides of the equation) and, moreover, trying to convert it to an ordinary eigenproblem would return an operator in the left member which would be no longer Hermitian. After having determined the eigensolution of (12.1), the corresponding electric field \vec{E} is found by applying to it the curl operator.

The eigenvalues $(\omega/c_o)^2$ are degenerate because, as stated by the Bloch-Floquet theorem [9, 16], the proper eigensolutions of (12.1) can be expressed as the superposition of *Bloch modes*:

$$\vec{H}_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} \vec{u}_{\vec{k}}(\vec{r})$$

where $\vec{u}_{\vec{k}}(\vec{r})$ is a periodic function on the lattice:

$$\vec{u}_{\vec{k}}(\vec{r} + \ell\vec{a}_1 + m\vec{a}_2 + n\vec{a}_3) = \vec{u}_{\vec{k}}(\vec{r})$$

($\ell, m, n = 0, \pm 1, \pm 2, \dots$) and \vec{k} is a vector ranging continuously, not in the whole dual k -space, but only in the (first) *Brillouin zone*. In fact in the k -space of (12.1) there is a *reciprocal* lattice defined by the reciprocal base vectors:

$$\begin{aligned} \vec{b}_1 &= 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \\ \vec{b}_2 &= 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \\ \vec{b}_3 &= 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \end{aligned}$$

satisfying

$$\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{ij},$$

which define a reciprocal unit cell. Labelling the Bloch modes with \vec{k} outside this reciprocal unit cell means redundancy. The first Brillouin zone is a finite volume region in the k -space around the origin such that one cannot get from one point of such volume to another by adding to the first any vector of the form:

$$\ell'\vec{b}_1 + m'\vec{b}_2 + n'\vec{b}_3$$

with $\ell', m', n' \in \mathbb{Z}$. This zone contains all the essential \vec{k} values: picking values outside it implies only the mere addition of an integer multiple of 2π to the imaginary exponential. The eigenproblem for $\vec{u}_{\vec{k}}(\vec{r})$ is found by inserting the Bloch mode in (12.1):

$$(i\vec{k} + \vec{\nabla}) \times \frac{1}{\epsilon(\vec{r})} (i\vec{k} + \vec{\nabla}) \times \vec{u}_{\vec{k}}(\vec{r}) = \left(\frac{\omega(\vec{k})}{c_o} \right)^2 \vec{u}_{\vec{k}}(\vec{r})$$

which has to be solved, for every \vec{k} fixed in the first Brillouin zone, with \vec{r} bounded in the unit cell defined by \vec{a}_1 , \vec{a}_2 and \vec{a}_3 . Due to this spatial domain bounding, the resulting spectrum of (angular) frequency eigenvalues is a discrete one, with an ordered sequence:

$$\omega_0(\vec{k}), \omega_1(\vec{k}), \omega_2(\vec{k}), \dots \quad (12.2)$$

which determines the *band structure* of the photonic crystal. If for two consecutive frequencies, ω_n and ω_{n+1} , it happens that a discontinuity is present between these two values at the boundary of the first Brillouin zone along a given \vec{k} direction, then a band gap has opened and electromagnetic radiation is not propagated in that direction, because it is not supported by the crystal material distribution.

In the time domain, the most general solution of the Maxwell's equations inside the crystal, considered as having infinite extension, is found by superposing Bloch modes:

$$\vec{H}(\vec{r}, t) = \sum_{n=0}^{\infty} \int_U A(\vec{k}) e^{i(\vec{k} \cdot \vec{r} - \omega_n(\vec{k})t)} \vec{u}_{\vec{k}}(\vec{r}) d^3\vec{k}$$

where A is an amplitude function determined by the initial conditions and U denotes the first Brillouin zone.

12.2 The FCC opal photonic crystal

For a face-centered cubic (FCC) crystal, the underlying spatial periodic structure is that resulting from the corner points of cubic cells of edge size d , but also with extra sites at the center of the faces. More precisely, the primitive lattice vectors, expressed in the usual base \hat{x} , \hat{y} , \hat{z} of a standard orthogonal cartesian reference frame are:

$$\vec{a}_1 = \frac{d}{2}(\hat{x} + \hat{y})$$

$$\vec{a}_2 = \frac{d}{2}(\hat{z} + \hat{x})$$

$$\vec{a}_3 = \frac{d}{2}(\hat{y} + \hat{z}).$$

The FCC opal photonic crystal considered here, see [57,58], is a self-assembled structure of polystyrene dielectric spheres with their centers occupying the spatial locations defined by linear combinations of the above lattice vectors with integer coefficients:

$$\vec{r}_{\ell,m,n} = \ell\vec{a}_1 + m\vec{a}_2 + n\vec{a}_3$$

$(\ell, m, n = 0, \pm 1, \pm 2, \dots)$. Pictorially, an example of the resulting structure is the one depicted in Fig. 12.1:

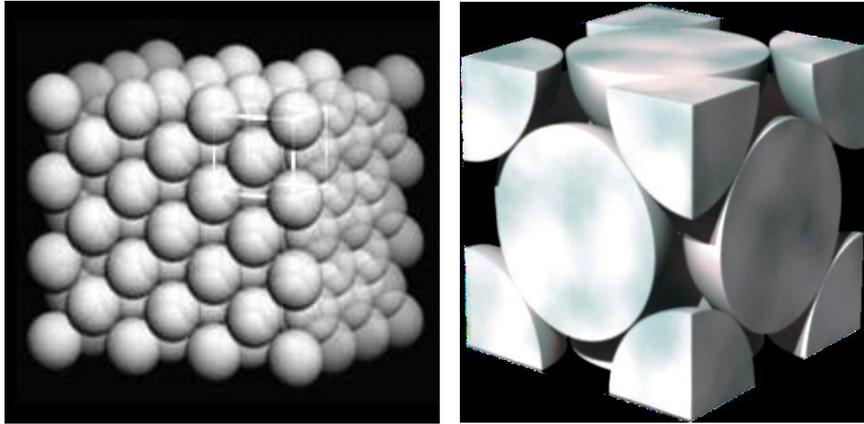


Figure 12.1: Example of an FCC lattice whose reticular positions are occupied by spherical objects.

The first Brillouin zone of this FCC lattice is represented by the truncated (regular) octahedron of Fig. 12.2:

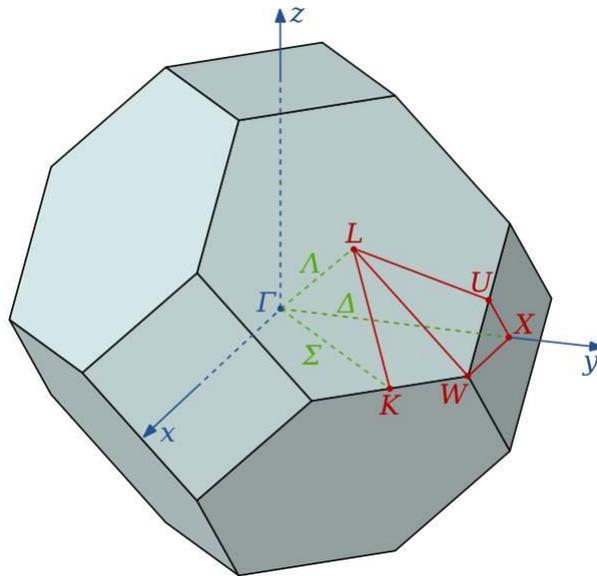


Figure 12.2: Truncated octahedron representing the first Brillouin zone of the FCC crystal lattice. The frame axes correspond to the x , y and z components of \vec{k} and should be better labelled as k_x , k_y and k_z .

in which Γ is the origin of the k -space and K, L, U, W, X denote typical points on its boundary, while Δ, Λ, Σ denote typical paths inside it. If:

$$\beta = \frac{2\pi}{d},$$

then the position vectors of such points are (using an orthonormal basis in the k -space):

$$\Gamma \equiv (0, 0, 0)$$

$$K \equiv \left(\frac{3}{4}\beta, \frac{3}{4}\beta, 0\right)$$

$$L \equiv \left(\frac{1}{2}\beta, \frac{1}{2}\beta, \frac{1}{2}\beta\right)$$

$$U \equiv \left(\frac{1}{4}\beta, \beta, \frac{1}{4}\beta\right)$$

$$X \equiv (0, \beta, 0)$$

$$W \equiv \left(\frac{1}{2}\beta, \beta, 0\right).$$

The radius R of the assembled spheres is estimated to be $R = 118$. nm, with a “thermodynamically resulting” FCC spacing of $d = 333.75$ nm, which corresponds to the maximal *filling factor*, the ratio of the void to the filled space, of 0.74.

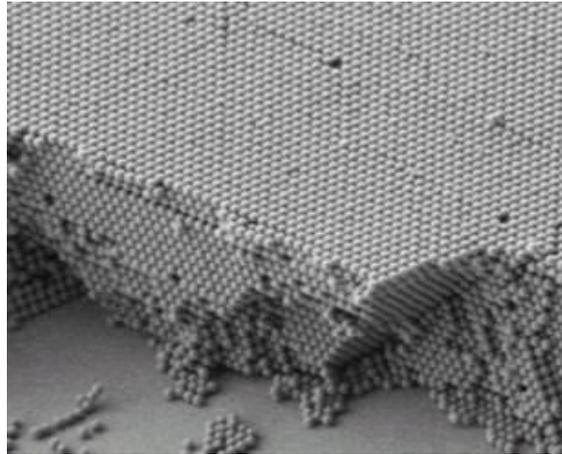


Figure 12.3: Scanning Electron Micrograph (SEM) image of a sample of the opal photonic crystal [57,58].

Polystyrene is a dielectric material with an assumed constant relative electric permittivity $\epsilon_r = 2.4$ on the given wavelength range without any dissipation,

i.e., with electric conductivity $\sigma = 0$ Siemens/m (1 Siemens = $1 \Omega^{-1}$). A microscopic view of the sample is given in Fig. 12.3.

Effectively, the crystal sample considered here is made of a stacked sequence of $(1, 1, 1)$ reticular planes, where Miller indices notation [16] is used to denote them. When this sample is optically analyzed, a beam of light impinges orthogonally onto this system of planes. Referring to Fig. 12.2 this direction corresponds, in the k -space, to the Λ path joining the origin Γ with the point L . To get more insight about the disposition of such $(1, 1, 1)$ planes with respect to the FCC lattice, one can refer to Fig. 12.4:

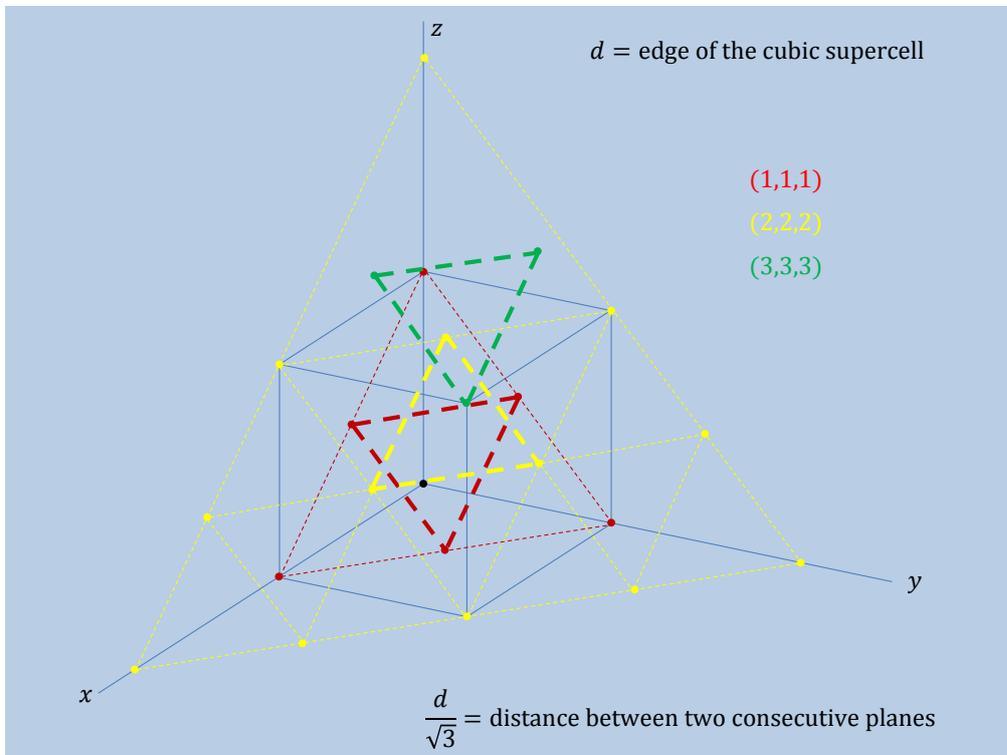


Figure 12.4: Figure showing how $(1, 1, 1)$ planes stems out from the FCC periodic structure.

The FCC structure viewed from the Λ direction is such that the stacked planes, the distance between each pair of them being $D = \frac{d}{\sqrt{3}}$, form a recursive sequence $ABCABC \dots$, and on each plane the lattice sites lie at the vertices of equilateral triangles of side length $b = \frac{d}{\sqrt{2}}$. See Fig. 12.5:

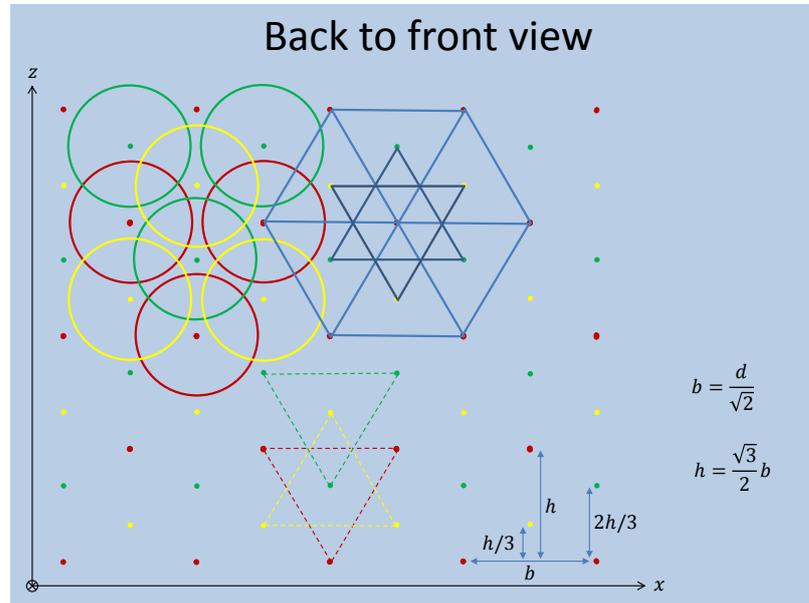


Figure 12.5: Sectional view showing the relative displacements among the three triangular tilings in the stacked A (red), B (yellow) and C (green) planes, as seen by the impinging light beam.

In passing from a A plane to the B and then to the C planes of the recursive sequence, the triangular tilings are shifted one with respect to the other as indicated by Fig. 12.5.

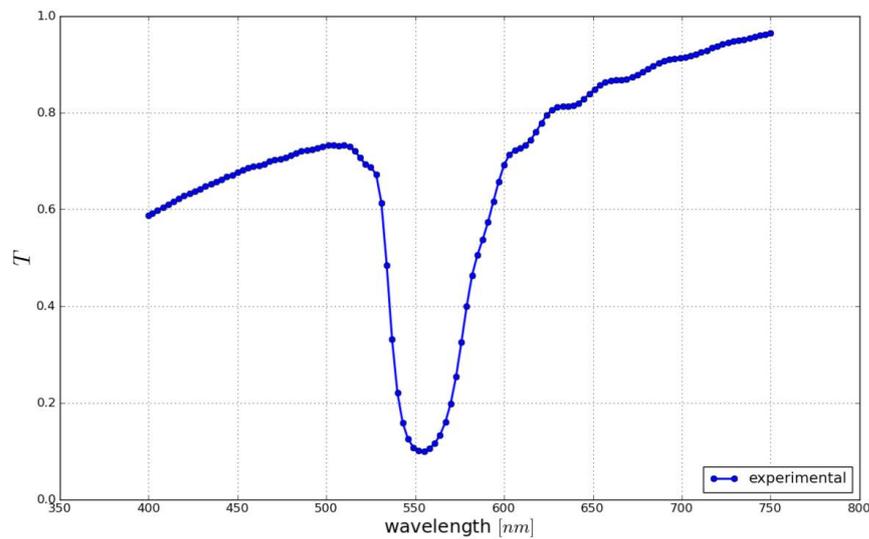


Figure 12.6: Experimental transmittance of a FCC crystal sample made of stacked $(1, 1, 1)$ planes. Thickness $t \approx \text{mm}$. Normal incidence.

When on a macroscopic sample slab of such an opal photonic crystal, of thickness $t \approx 1$ mm, impinges normally a light beam, the laboratory measured transmittance is the one shown in Fig. 12.6 [57]. Clearly, there is a band gap near 550 nm due to the particular structure of the crystal which, acting as a Bragg reflector [9], prevents light propagation. This effect is a purely classical one and can be reproduced by an FDTD numerical modelization of the type described in Part II of the present work and which is described schematically in Fig. 12.7:

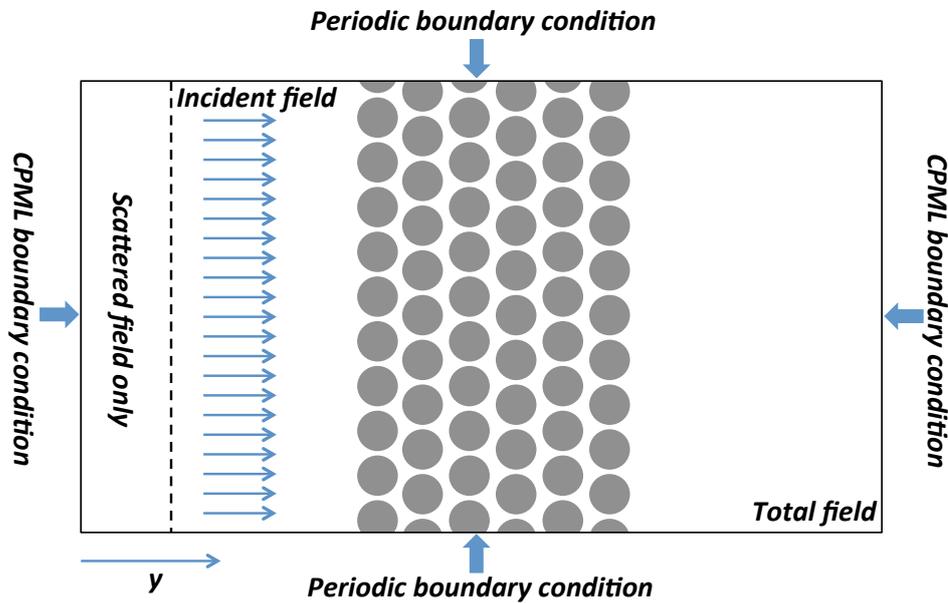


Figure 12.7: 2D schematic of the FCC opal photonic crystal FDTD modeling.

The computational domain truncation is got by means of CPML absorbing boundary conditions, described in Section 6.1. However, these are applied only to the upstream and downstream external faces, with respect of the impinging beam propagation direction along the y axis. Periodic boundary conditions (PCB), described in Section 6.2, are applied on the surfaces with normals perpendicular to the beam propagation direction, i.e., the outer xy and yz planes, to simulate an ideal crystal slab of infinite lateral extension. They consist in reinjecting the field values in a wraparound fashion. To this end, the interacting crystal structure has to be represented in a consistent manner: the part left out at one side must be mirrored at the other opposite side [59]. In Fig. 12.8:

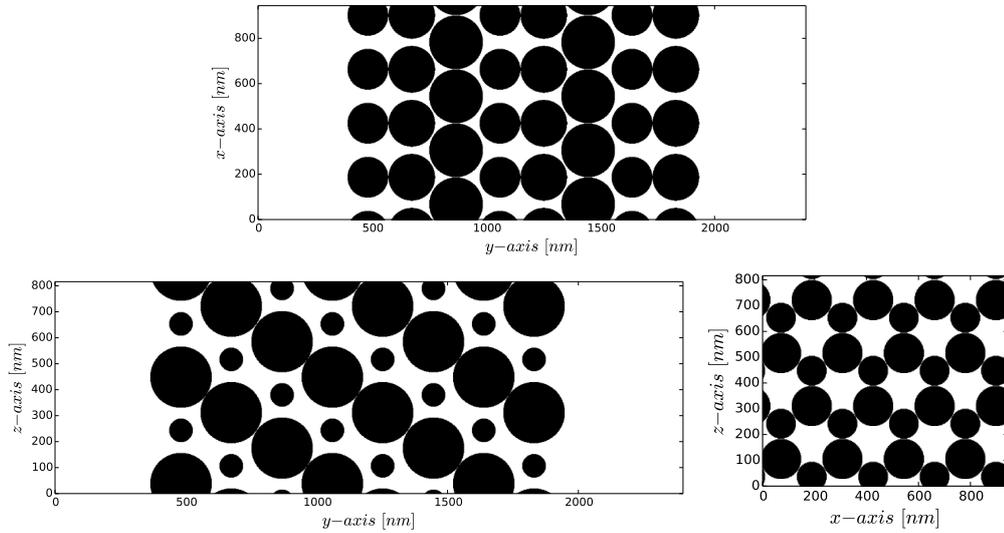


Figure 12.8: Lateral xy (top), horizontal yz (bottom-left) and front xz (bottom-right) sections of the FCC lattice as modeled in the FDTD code.

are shown sectional views depicting how the polystyrene spheres are packed in the stack of $(1, 1, 1)$ planes along the y axis, to recreate the FCC lattice of particles, as they result after the model structure generation of the FDTD code has been applied. The numerical simulations used cubic Yee cells of edge size $\delta_{x,y,z} = 2$ nm, with a total of $X \times Y \times Z$ Yee cells along the x , y and z axes respectively, which took x seconds for x iterations on x parallel MPI processes of a BlueGene/Q supercomputer. A pulse like the ones of Fig. 5.6 was “injected” as described in Section 7.1, but with no lateral and ending guiding surfaces, starting from the xz surface corresponding to the vertical dashed line of Fig. 12.7. Behind that surface only the reflected field is present. In that region, if sufficiently far from the crystal slab, reflectance is evaluated by integrating the real part of the normal component of the complex Poynting vector. The transmittance is similarly evaluated by considering downstream planes beyond the crystal slab in the total field region (incident plus transmitted ones). The FDTD numerically calculated reflectance/transmittance result for a stack of 20 $(1, 1, 1)$ planes is given in Fig. 12.9 [60]; the transmittance T and reflectance R curves in the graph are calculated independently, by integrating the calculated Poynting vector upstream in the scattered field region and downstream in front of the slab respectively, and not by using the relationship $R + T = 1$, which holds without absorption. It is apparent a bandgap that opens around 540 nm. In Fig. 12.10 it is shown the FDTD numerically calculated transmittance T result for a stack of 100 $(1, 1, 1)$ planes, as compared with the experimentally measured curve already shown in Fig. 12.6 [60].

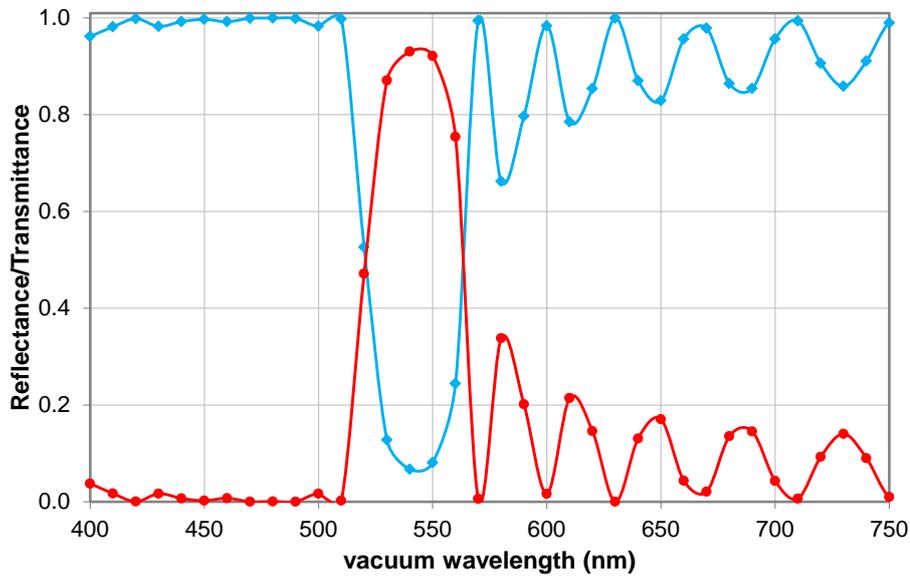


Figure 12.9: Reflectance and transmittance of an FDTD numerical FCC opal photonic crystal made of 20 (1, 1, 1) stacked planes, independently calculated to test the $R + T = 1$ relationship validity, for inherent coherence of the numerical model probing.

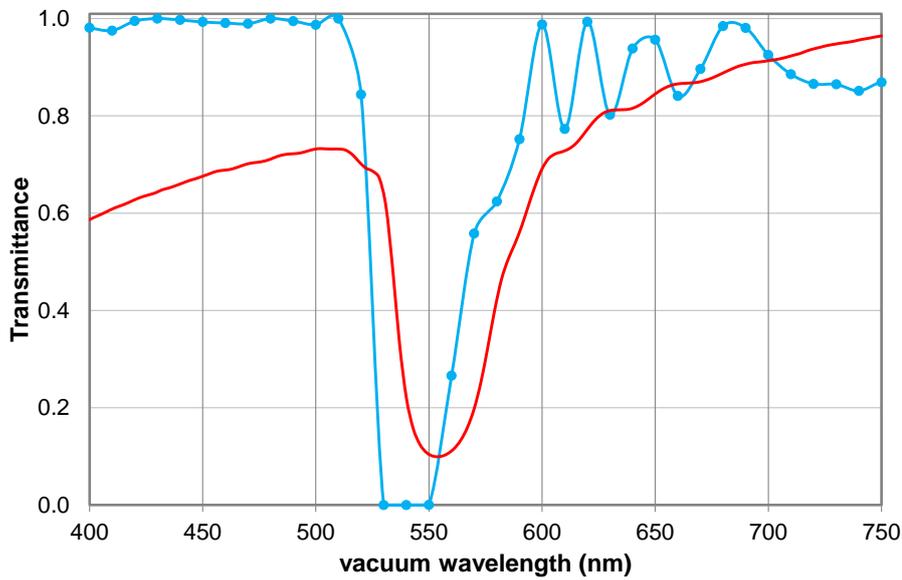


Figure 12.10: Transmittance of an FDTD numerical FCC opal photonic crystal made of 100 (1, 1, 1) stacked planes, as compared with the one of a real macroscopic crystal sample having the same structure.

There is a fairly good agreement in determining the band gap near 550 nm, demonstrating the effectiveness of the numerical FDTD modeling in describ-

ing the physical behavior of such photonic structures. One should keep in mind that R and T are evaluated at a discrete set of (angular) frequencies of analysis, after the fields for the power flux evaluation on the given surfaces, by means of the Poynting vector, have been Fourier transformed (by means of a DFT) from the time domain ones in the course of the simulations. Outside the band gap, the numerical reflectance and transmittance exhibit realistic oscillations which do not compare in the experimental curve, due to a smoothing process introduced by the measuring instrument. Should the frequency resolution be increased, and up to the resolution limitation exposed in Section 5.3, these oscillations would appear more pronounced. The color maps that follows show the electric field distribution inside the crystal sample, with 20 stacked $(1, 1, 1)$ planes, in correspondence of given planes parallel to the coordinate ones, extracted from the computational volume bulk at the three vacuum wavelengths of 500 nm, 550 nm and 600 nm respectively [60]. The incident field is linearly polarized along the z axis, i.e., normal to the plane of figures in Fig. 12.11 but along the vertical axis in Fig. 12.12:

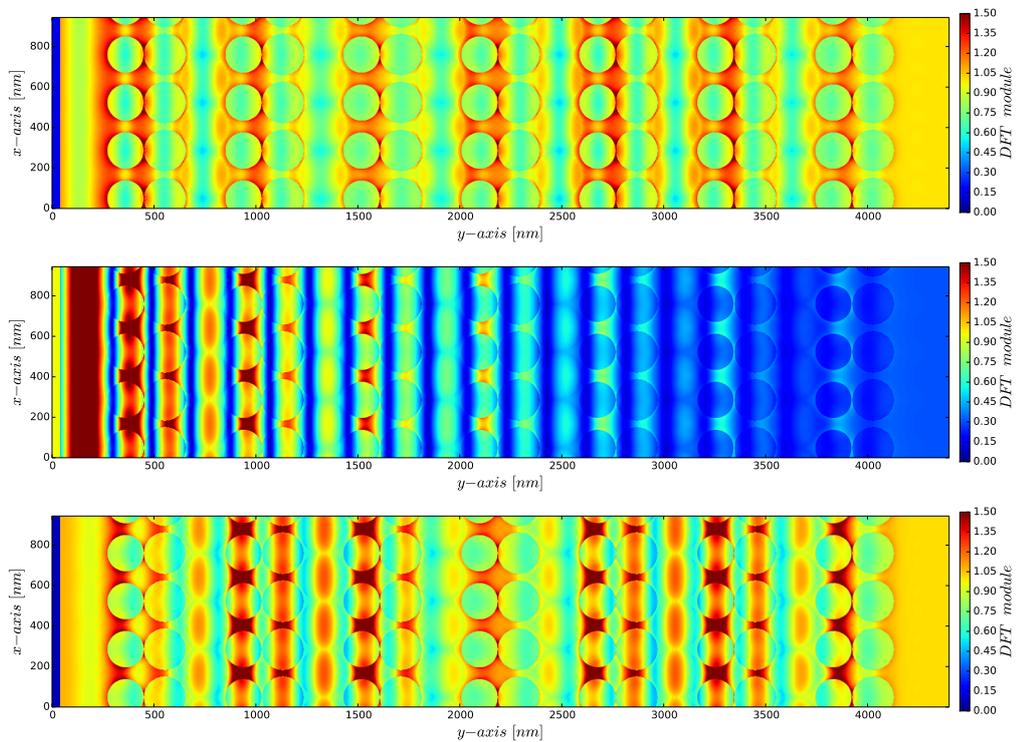


Figure 12.11: Normalized DFT module of \vec{E} on a xy plane inside the crystal sample at 500 nm (top), 550 nm (middle) and 600 nm (bottom).

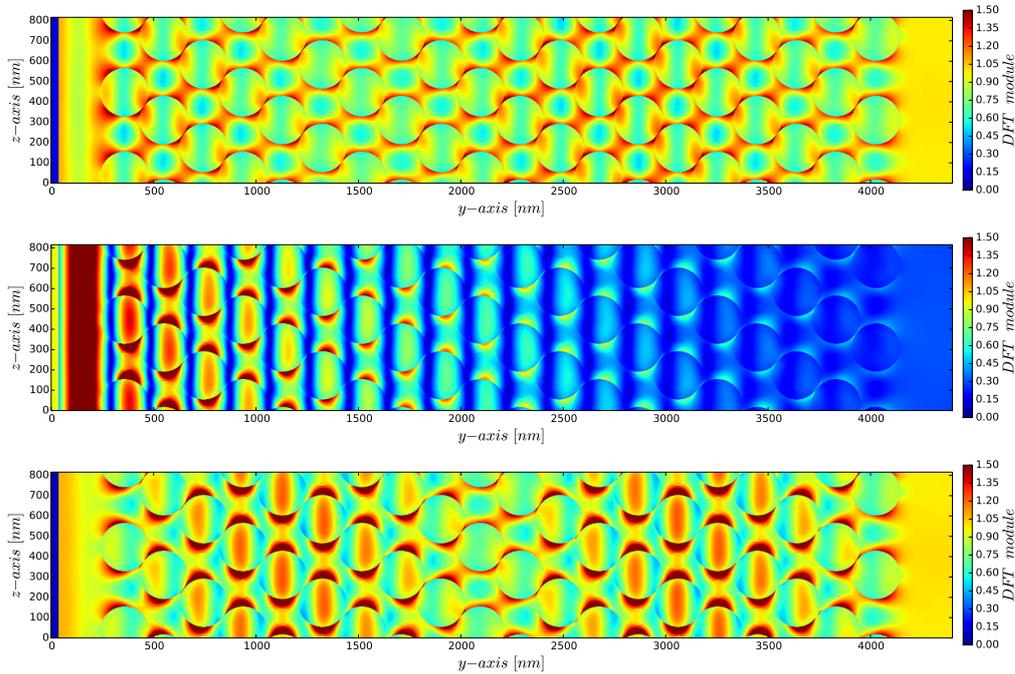


Figure 12.12: Normalized DFT module of \vec{E} on a yz plane inside the crystal sample at 500 nm (top), 550 nm (middle) and 600 nm (bottom).

as can be noted, and coherently with the transmittance band gap, the 550 nm field has much less penetration inside the crystal than wavelengths outside the band gap do. This is an evidence that, at the band gap frequency, multiple reflections from the $(1, 1, 1)$ lattice planes sum up with phase coherency in such a way to prevent any electromagnetic energy transport down the crystal along the propagation y axis.

Earlier simulations [59] made without the use of the periodic boundary conditions, but considering a finite crystal sample in all the three directions, gave less clear results than those above, due to edge effects of the finite sample and to the low dielectric contrast of the spheres. In particular, the bandgap was reproduced with a very small depth and to get a well indented transmittance spectrum a lot of $(1, 1, 1)$ lattice planes would have to be stacked in the numerical model, similarly to what happens in the experimental situation in which a macroscopic crystal sample is used. Actually, it is likely that secondary irradiation from the lateral sample surfaces bends the planar wavefronts thus preventing a full Bragg destructive interference in the forward direction. However, increasing the number of modeled reticular planes requires a lot of computational resources which can be spared through the use of the periodic boundary conditions. In any case, with this setup using a finite crystal sample, there is evidence of the band gap presence by analyzing

the scattered far field radiation in the solid angle around the crystal target by means of a numerical implementation the Kirchoff integral formula explained in Section 2.3, and which uses the numerically FDTD calculated near electric and magnetic field values around the crystal. Examples of such numerically calculated radiation angular distribution are given in Fig. 12.13:

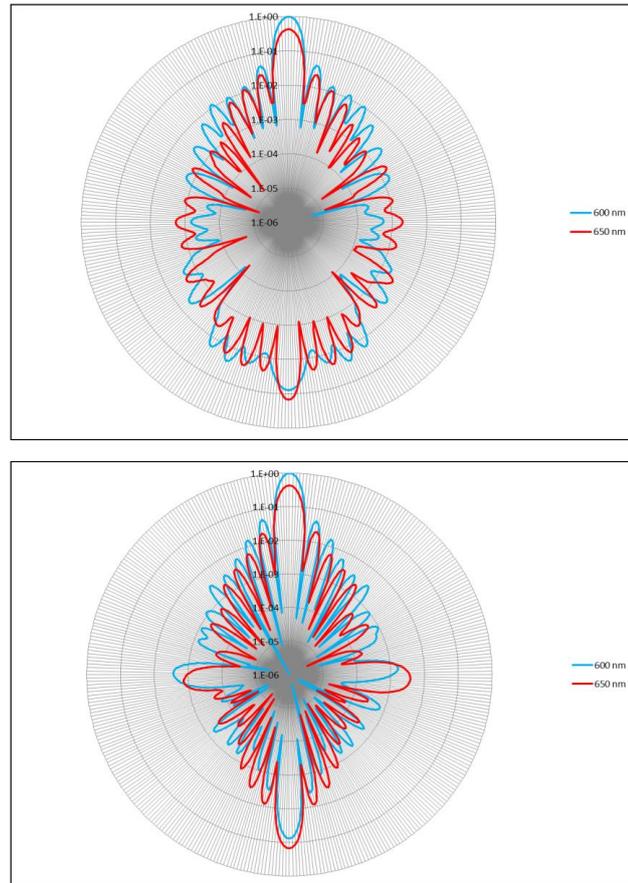


Figure 12.13: .

The transmittance in this case is evaluated as:

$$\mathcal{T} = 1 - \frac{P_r}{P_i}$$

where P_r is the total radiated power and P_i is the power incident of the crystal target of finite sizes, for a beam cross section greater than that of the sample.

12.3 Multilayer equivalence: homogenization

The discontinuous system of stacked $(1, 1, 1)$ planes above described, each one consisting in a triangular tiling of spherical dielectric particles, can ultimately be reduced, by an heuristic procedure here described, to a material having an electric permittivity varying, in a periodic fashion, only in the y impinging direction. This material can be thought as a finely periodic layered medium with a step (or staircase) ε function, whose transmittance/reflectance properties can be analyzed by means of the techniques exposed in Chapter 4. An array of small thickness, *homogeneous*, stacked slabs can give raise to high reflectance as the result of the in-phase superposition of a large number of weak reflections. This resonances would correspond to band gaps (forbidden bands) in the crystal for which the Bloch modes become evanescent and do not propagate down the crystal itself. The light energy is thus totally reflected and the array becomes a Bragg reflector. The heuristic homogenization procedure is illustrated schematically by Fig. 12.14:

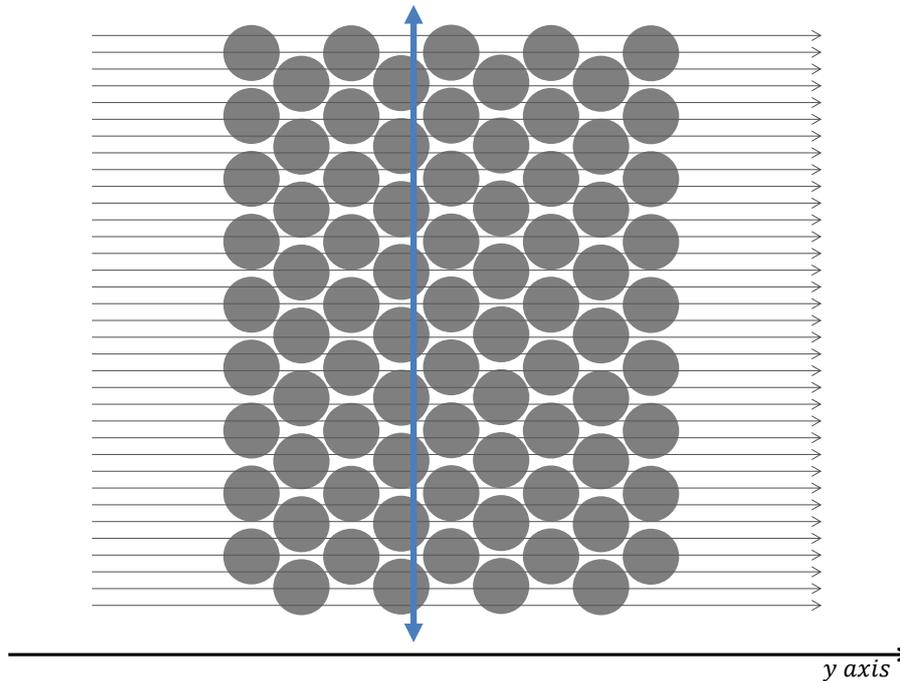


Figure 12.14: 2D schematic view of the homogenization procedure with the probing plane (blue) shifting perpendicularly to the stacked system of $(1, 1, 1)$ planes. Also shown is the pencil of perpendicular lines intercepting the system of dielectric spheres.

A sliding probing plane is moved discretely along the crystal longitudinal y

axis (which is also the beam propagation axis). At each position, the intersections of that plane with the system of spheres of which the crystal is made are determined. A pencil of straight lines is projected along the y axis. The starting points of these lines are uniformly distributed in the xz plane normal to the y axis. The points of intersections of the lines with the probing plane are also determined. To those points which also belong to the intersections with the spheres, a polystyrene ε value is assigned. To those which do not belong, a free space ε is assigned. An averaging procedure of such points is made for the whole probing plane at that specific y position, which allows to define an $\varepsilon(y)$ function value. By moving the plane at different positions along the y axis, a profile of the refractive index $n = \sqrt{\varepsilon_{avg}/\varepsilon_o}$ is obtained like the one shown in Fig. 12.15:

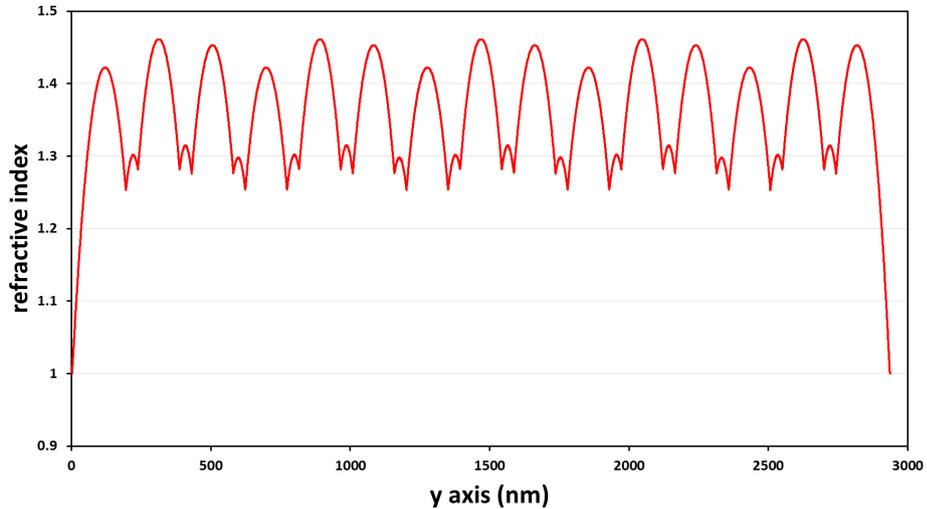


Figure 12.15: Refractive index of an FCC opal crystal sample after homogenization.

This homogenization has been got by sampling with 8000 points the longitudinal y axis of a 3000 nm length crystal made of 15 stacked $(1, 1, 1)$ planes. This corresponds to a sampling step $\delta y = 0.36721$ nm. It is apparent a periodically repeating pattern, made of three successive big peaks (including the two intermediate small peaks) plus a nearby small peak. They correspond to the ABC sequence of reticular $(1, 1, 1)$ planes. This repeating pattern occupies 1574 consecutive sampling points along the y axis. To get the averaged value at each y in the graph, a pencil of 600×600 straight lines uniformly distributed over an area corresponding to about 5×5 dielectric spheres in the xz plane has been considered. One recognizes that, when the sampling is made finer and finer, the profile reaches a limit pattern and no longer changes from a certain sampling refinement onward. Then a sequence of 1574 con-

secutive slabs of thickness 0.36721 nm each, repeated 100 times, has been considered as a periodic multilayered medium. To this system, the transmittance matrix method of Chapter 4 has been applied over the wavelength range $400 \div 800$ nm. The matrix method of Chapter 4 is an exact analytical approach that generalizes the single interface Fresnel's coefficients concept. It is implemented in a computer code because it is applied for calculating a spectrum of a very large number of nearby slabs. The resulting \mathcal{T} is as shown in Fig. 12.16:

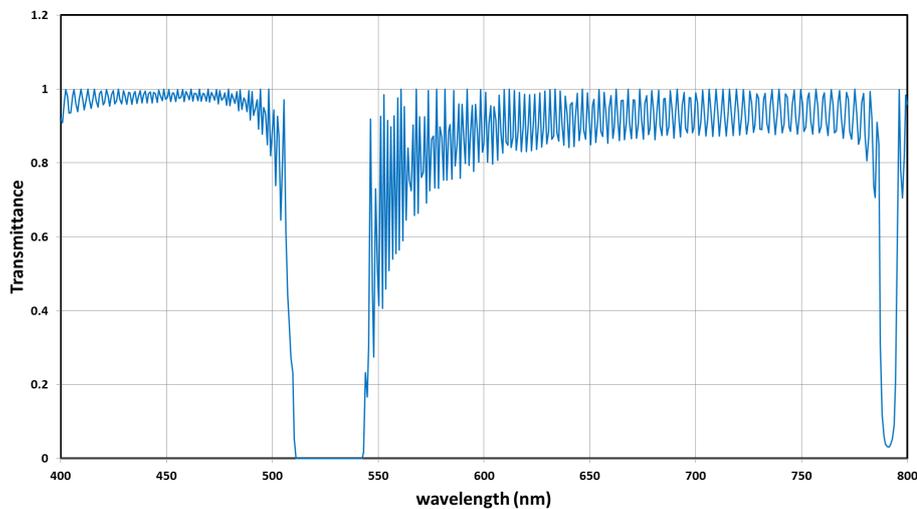


Figure 12.16: .

which is in fairly good accordance with both the experiment and the FDTD numerical method: see Fig. 12.10. It also predicts a narrower band gap at a frequency corresponding to the vacuum wavelength of about 790 nm, which is outside the wavelength range considered so far. The graph in Fig. 12.16 is also more detailed in describing the oscillations outside the band gaps, because it is based on an exact analytical calculation.

The homogenization procedure here described is very promising in offering an insight into the crystal transmittance/reflectance behavior, before starting any large FDTD simulations which only, in any case, can give an in-depth analysis of the electromagnetic field inside and near the crystal itself. One of the major further advancements of this homogenization procedure, would be one able to deduce the *tensorial* character of the averaged ε , when applied to opal crystals including particles with less symmetry than the spherical ones of the present case.

12.4 Dispersion band structure reconstruction

By a post-processing phase, once an FDTD simulation has been performed on an opal crystal sample of given structure — in the present case, the FCC one — and the DFT complex electric field data inside the crystal are allowable at a discrete set of (angular) frequency values, it is possible to proceed at a crystal band structure (12.2) reconstruction, for \vec{k} ranging in its first Brillouin zone, like the one depicted in Fig. 12.2. In the present case only the \vec{k} values belonging to the Λ path will be analyzed [60] which, corresponding to the direction of incidence of the primary field of the beam of light, will label the modes most easily excited inside the crystal bulk. First of all, a multilayer analysis like the one described in the previous Section 12.3, allows to identify the band gaps at a glance from the transmittance spectrum. Then, for each discrete angular frequency FDTD analyzed, below and above the band gap, the corresponding spatial distributions of the electric DFT data are subjected to a spatial Fourier transform, using:

$$\iiint \tilde{E}_z(\vec{r}; \omega) e^{-i\vec{k}\cdot\vec{r}} d^3\vec{r} \approx \sum_{m,n,p} \tilde{E}_z(m, n, p; \omega) e^{-i\vec{k}\cdot\vec{r}_{m,n,p}}$$

where $\vec{r}_{m,n,p}$ denotes discrete Yee FDTD grid positions and $\tilde{}$ denotes frequency transformed field values. In the present case only the z component along the direction of polarization of the incident field has been considered, although components along other cartesian axes are also excited within the crystal, but with considerable less intensity. In the above expression $k = \|\vec{k}\|$ ranges discretely from 0 to $\sqrt{3}\beta$ (two times the length of the ΓL segment in Fig. 12.2). As a vector, \vec{k} is kept fixed in the \hat{y} direction, because the crystal is already rotated in such a way that Γ coincides with the y beam propagation axis. The necessarily finite integration domain is positioned at the center of the crystal sample, embracing some FCC unit cells along the beam propagation direction, but with the boundaries well apart from those of the sample, to avoid contributions from unwanted reflected components. This implies some negligible truncation errors in the tails of the calculated spectra. The integration domain may instead be extended along the directions perpendicular to the faces at which periodic boundary conditions are imposed, until they are met, without any truncation effects. The Bloch mode supported for a given angular frequency ω is identified from the corresponding k -spectrum as a dominant peak, see Fig. 12.17:

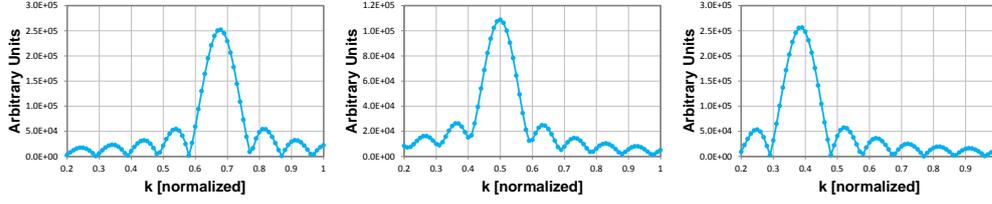


Figure 12.17: Examples of numerically calculated spatial spectra for $\omega = 2\pi c/\lambda$, with $\lambda = 400$ nm (left), 540 nm (middle) and 700 nm (right). Normalization of the horizontal axis is $\frac{k}{\sqrt{3}\beta}$.

which is a normalization independent operation. What we are really interested in is not the peak height, but the k value of the peak occurrence. Such peak detection has been made finer by fitting the points around the identified peak through a spline interpolation, to better localize the value at the maximum. By comparing the relative height of the peaks allows us also to recognize the bandgap along the abscissas k -axis, because a non-propagating mode has a lower amplitude spectrum, compared with those of the propagating ones. On the other hand, the (angular) frequency ω occurrence of the bandgap is also already known from the transmittance curve, like the ones in Fig. 12.9 or Fig. 12.10. There, the abscissas are labelled as vacuum wavelengths for more practical insight, but actually correspond univocally to frequency domain data: $\omega = 2\pi c/\lambda$. Fig. 12.18:

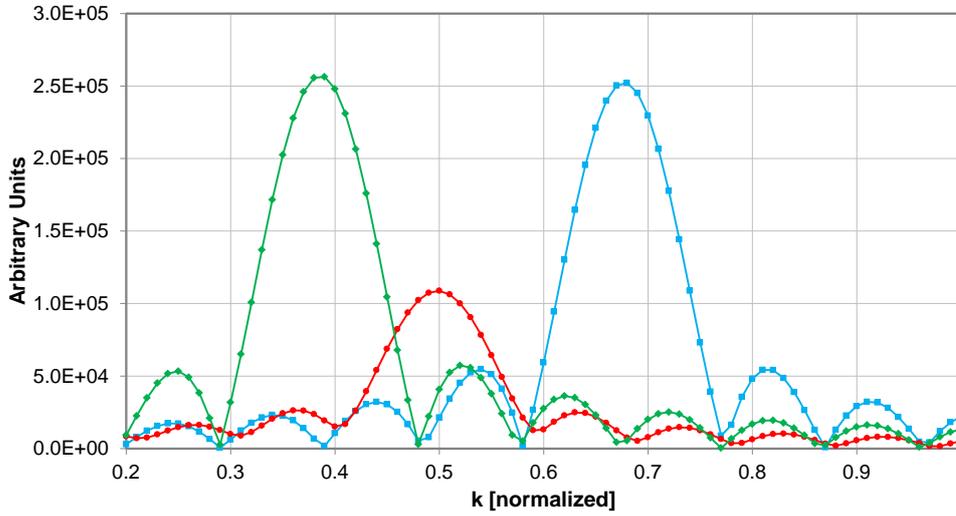


Figure 12.18: Numerical *spatial* amplitude spatial spectra of $\tilde{E}_z(\vec{r}; \omega)$ for $\omega = 2\pi c/\lambda$, with $\lambda = 700$ nm (left curve), 540 nm (middle curve) and 400 nm (right curve). Normalization of the horizontal axis is $\frac{k}{\sqrt{3}\beta}$.

shows together the graphs of the previous one. Note that, as expected, at the angular frequency corresponding to $\lambda = 540$ nm (the middle graph), i.e.

the frequency of the first bandgap (see Fig. 12.9), the magnitude of the peak is less than one half the magnitudes of the other peaks (left and right graphs in the figure) which correspond to fully propagated modes.

One eventually ends up with a sequence of (k, ω) pairs which, when graphed, allows the quantitative reconstruction of the dispersion curve searched for. That in Fig. 12.19 is for a stack of 20 reticular $(1, 1, 1)$ planes, in a sequence $ABCABC \dots$. The k and ω axes of the graph are normalized: the horizontal one corresponds to $k/\beta\sqrt{3}$ values, the vertical one to $\omega d/2\pi c$ values. To get the second branch of the curve above the first bandgap (the so called “air-line”), the relevant k data obtained by means of the spatial Fourier transforms have been mirrored in the vertical line through the value $k_{bg} = 0.499\beta$ corresponding to the first bandgap. It is apparent from the figure that the slope flattens at the interval extrema on the k -axis — except at the origin — as expected. The four dots standing at the most right position in the graph of Fig. 12.19 correspond to the bandgap and should be neglected. The width of the bandgap here corresponds to that of Fig. 12.9.

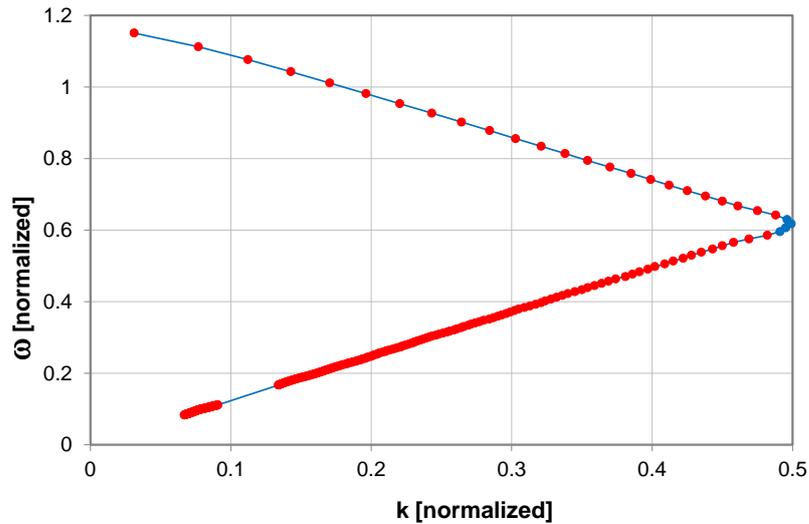


Figure 12.19: Quantitative reconstruction of the dispersion curve of an opal FCC photonic crystal.

The behavior of the curve toward the origin is, from a linear regression analysis, that of a straight line with zero intercept. The slope of the straight line allows to calculate the effective refractive index of the crystal. One gets a numerically predicted value of $n_{eff} = 1.393$; the experimental value of $n_{eff} = 1.436$ reported by [58] and that of $n_{eff} = 1.427$ from geometric mean consideration in effective medium approximation there cited are in fairly good accordance with ours.

That described is a general procedure, usable for different packing geome-

tries and which may be also extended to the higher branches of the dispersion curve and for the entire Brillouin zone. One major difficulty with this generalization is to achieve the suitable orientation of the impinging beam according to the direction in the k -space of the crystal which one wants to analyze, in such a way to get a “maximal excitation” of the involved Bloch modes. To this end it is clear that, instead to rotate the crystal, it is better to give a general propagation direction to the beam of light, as described in Section 7.2. Nonetheless, it remains open the problem of what are the periodic boundary conditions to be used in this case of non normal incidence, because an unknown phase difference will arise on the opposite outer faces of the computational volume.

Appendix A

Some dyadics calculations

All greek letters used as indices in the sums that follows take the values x, y and z . For a dyad $\vec{a}\vec{b}$, the vector product from the left with the vector \vec{c} is:

$$\vec{c} \times (\vec{a}\vec{b}) \stackrel{\text{def}}{=} (\vec{c} \times \vec{a})\vec{b}$$

(multiplying from the right would produce a different result). From this definition follows the one for the curl of a dyadic given in Chapter 2 and below reproduced:

$$\vec{\nabla} \times \vec{G} = \sum_{\alpha} \sum_{\beta} \sum_{\gamma} \partial_{\gamma} G_{\alpha\beta} (\hat{\gamma} \times \hat{\alpha}) \hat{\beta}.$$

By introducing the Levi-Civita antisymmetric symbol:

$$\varepsilon_{\alpha\beta\gamma} = \begin{cases} +1 & \text{if } (\alpha, \beta, \gamma) = (x, y, z), (y, z, x), (z, x, y) \\ -1 & \text{if } (\alpha, \beta, \gamma) = (z, y, x), (x, z, y), (y, x, z) \\ 0 & \text{if } \alpha = \beta \text{ or } \beta = \gamma \text{ or } \gamma = \alpha \end{cases},$$

one can write the above expression like:

$$\vec{\nabla} \times \vec{G} = \sum_{\alpha} \sum_{\beta} \sum_{\gamma} \sum_{\tau} (\partial_{\gamma} G_{\alpha\beta}) \varepsilon_{\tau\gamma\alpha} \hat{\tau} \hat{\beta}.$$

Taking the curl again:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{G} = \sum_{\alpha} \sum_{\beta} \sum_{\gamma} \sum_{\tau} \sum_{\sigma} (\partial_{\sigma} \partial_{\gamma} G_{\alpha\beta}) \varepsilon_{\tau\gamma\alpha} (\hat{\sigma} \times \hat{\tau}) \hat{\beta}$$

or:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{G} = \sum_{\alpha} \sum_{\beta} \sum_{\gamma} \sum_{\tau} \sum_{\sigma} \sum_{\rho} (\partial_{\sigma} \partial_{\gamma} G_{\alpha\beta}) \varepsilon_{\tau\gamma\alpha} \varepsilon_{\rho\sigma\tau} \hat{\rho} \hat{\beta}.$$

But:

$$\sum_{\tau} \varepsilon_{\tau\gamma\alpha} \varepsilon_{\rho\sigma\tau} = \sum_{\tau} \varepsilon_{\tau\gamma\alpha} \varepsilon_{\tau\rho\sigma} = \delta_{\gamma\rho} \delta_{\alpha\sigma} - \delta_{\gamma\sigma} \delta_{\alpha\rho}.$$

After insertion in the expression for $\vec{\nabla} \times \vec{\nabla} \times \vec{G}$ there remains:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{G} = \sum_{\alpha} \sum_{\beta} \left[\left(\sum_{\gamma} \partial_{\alpha} \partial_{\gamma} G_{\beta\gamma} \right) - \left(\sum_{\gamma} \partial_{\gamma}^2 G_{\alpha\beta} \right) \right] \hat{\alpha} \hat{\beta}$$

where use has also been made of the symmetry $G_{\gamma,\beta} = G_{\beta,\gamma}$. Remembering the expression for the components of \vec{G} in terms of g given in Chapter 2 one eventually gets:

$$\vec{\nabla} \times \vec{\nabla} \times \vec{G} = \sum_{\alpha} \sum_{\beta} (\partial_{\alpha} \partial_{\beta} g) \hat{\alpha} \hat{\beta} - \sum_{\alpha} \left(\sum_{\gamma} \partial_{\gamma}^2 g \right) \hat{\alpha} \hat{\alpha}$$

which is the searched result because the first term simplifies with the $\vec{\nabla}' \vec{\nabla}'$ part in the $k^2 \vec{G}$ term, while g is the fundamental solution of the scalar Helmholtz equation.

To demonstrate the dyadic divergence identity (2.15) it suffices to consider a dyad $\vec{A} \vec{B}$ in place of the dyadic \vec{D} , because the general case is a linear combination of dyads:

$$\begin{aligned} \vec{\nabla} \cdot (\vec{F} \times \vec{A} \vec{B}) &= \vec{\nabla} \cdot \left[(\vec{F} \times \vec{A}) \vec{B} \right] = \left[\vec{\nabla} \cdot (\vec{F} \times \vec{A}) \right] \vec{B} = \\ &= (\vec{A} \cdot \vec{\nabla} \times \vec{F} - \vec{F} \cdot \vec{\nabla} \times \vec{A}) \vec{B} = (\vec{\nabla} \times \vec{F}) \cdot \vec{A} \vec{B} - \vec{F} \cdot \vec{\nabla} \times (\vec{A} \vec{B}) \end{aligned}$$

because an operator acting from the left on a dyad acts only on the left vector. By replacing back $\vec{A} \vec{B}$ with \vec{D} one has (2.15).

Appendix B

Some Gaussian integrals

The starting point is represented by the classical definite integral:

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2} dx = 2 \int_0^{+\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}} \quad (\text{B.1})$$

which holds for a real $\alpha > 0$. This integral can be *analytically continued* for $\alpha \in \mathbb{C}$ (the integrand still being an even function of x) with $\text{Re}\{\alpha\} > 0$, by considering the analytic function:

$$e^{-\alpha z^2}$$

in the complex z -plane. Writing α in polar form:

$$\alpha = |\alpha|e^{i\theta}$$

with $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ one can consider the contour integration in the complex plane along the right triangular path with a vertex at the origin, a cathetus made of the interval $[0, \xi]$ of the real axis, the hypotenuse made of the half-line:

$$z = te^{-i\frac{\theta}{2}}$$

with the parameter $t \geq 0$, from which:

$$dz = e^{-i\frac{\theta}{2}} dt,$$

and by taking the remaining cathetus, which closes the path, as a vertical segment parallel to the imaginary axis. By letting $\xi \rightarrow +\infty$, this latter side contributes nothing to the integral because there is an exponentially damping factor with ξ and, having the integrand no singularities in the region enclosed by the path, the searched analytical continuation result follows.

A further generalization of (B.1) is the inclusion of a term linear in the variable of integration in the exponential:

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2 + i\beta x} dx = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}} \quad (\text{B.2})$$

with $\alpha, \beta \in \mathbb{C}$ and $\text{Re}\{\alpha\} > 0$. The above result follows after completing the square and shifting the origin of the x axis. One eventually is led to the integral:

$$\int_{-\infty}^{+\infty} e^{-\alpha(x-i\kappa)^2} dx$$

where κ is a real parameter. By integrating $e^{-\alpha z^2}$ on a rectangular path with horizontal sides going from $-\infty$ to $+\infty$, one on the real axis and the other passing through the $i\kappa$ point of the imaginary axis, being vanishing the contribution from the two vertical sides, one sees that the value of the above integral is the same as (B.1), thus confirming the result in (B.2).

Appendix C

Stability condition for the FDTD method

The stability criterion for the completely explicit FDTD bulk algorithm described in Chapter 5 is deduced here — following [40] — for an homogeneous nonabsorptive ($\sigma = 0$) source free medium of infinite extent with absolute parameters ϵ , μ and light speed $c = 1/\sqrt{\epsilon\mu}$. By introducing the temporally sampled composite field (with n the temporal index):

$$\vec{F}^n(\vec{r}) = \begin{bmatrix} \vec{E}^n(\vec{r}) \\ \vec{H}^{n+\frac{1}{2}}(\vec{r}) \end{bmatrix},$$

where account is taken of the half time step sampling shift between the electric and magnetic fields and by considering the inverse space Fourier transform (by an abuse of notation the same letters are used to denote the \vec{r} -domain or \vec{k} -domain field variables):

$$\vec{F}^n(\vec{r}) \propto \int \vec{F}^n(\vec{k}) e^{i\vec{k}\cdot\vec{r}} d^3\vec{k},$$

the time level transition corresponding to the finite-difference expressions (5.1), (5.2) can be written compactly in matrix form as:

$$\vec{F}^{n+1}(\vec{k}) = \begin{bmatrix} \hat{I} & \frac{\delta_t}{\mu} \hat{A} \\ -\frac{\delta_t}{\mu} \hat{A} & \hat{I} - \frac{\delta_t^2}{\epsilon\mu} \hat{A}^2 \end{bmatrix} \vec{F}^n(\vec{k}) = \hat{G}(\vec{k}) \vec{F}^n(\vec{k}).$$

$\hat{G}(\vec{k})$ is a 6×6 matrix operator containing as 3×3 submatrices the identity \hat{I} and:

$$\hat{A}(\vec{k}) = \begin{bmatrix} 0 & -2i \frac{\sin(k_z \delta_z/2)}{\delta_z} & 2i \frac{\sin(k_y \delta_y/2)}{\delta_y} \\ 2i \frac{\sin(k_z \delta_z/2)}{\delta_z} & 0 & -2i \frac{\sin(k_x \delta_x/2)}{\delta_x} \\ -2i \frac{\sin(k_y \delta_y/2)}{\delta_y} & 2i \frac{\sin(k_x \delta_x/2)}{\delta_x} & 0 \end{bmatrix}$$

where $\vec{k} = (k_x, k_y, k_z)$ and δ_x, δ_y and δ_z are the FDTD grid space steps. \hat{G} is the amplification or transition operator. Its norm has to be less than unity to prevent, by successive iterations, any fluctuation in the numerical solution to exponentially grow unbounded, thus overwhelming the true solution. By using the spectral norm of \hat{G} one is led to the eigenvalues of the matrix, which are 1 and:

$$\lambda_{\pm} = 1 - \frac{a^2}{2} \pm \frac{a}{2} \sqrt{a^2 - 4}$$

where

$$a = \frac{2\delta_t}{\sqrt{\epsilon\mu}} \sqrt{\frac{\sin^2(k_x \delta_x/2)}{\delta_x^2} + \frac{\sin^2(k_y \delta_y/2)}{\delta_y^2} + \frac{\sin^2(k_z \delta_z/2)}{\delta_z^2}},$$

all with double multiplicity. For the spectral radius to be not more than unity, it must be $a^2 \leq 4$, which happens if:

$$\frac{\delta_t}{\sqrt{\mu\epsilon}} \leq \frac{1}{\sqrt{\frac{1}{\delta_x^2} + \frac{1}{\delta_y^2} + \frac{1}{\delta_z^2}}},$$

which is the Courant-Friedrichs-Lewy (CFL) stability condition. It is customary to take $\delta_x, \delta_y, \delta_z$ equal to a common value δ . In this case the stability is:

$$\delta_t \leq \frac{\delta}{c\sqrt{3}}.$$

If there are more media with different phase velocities one has to use c_{max} in the upper bound limit for δ_t , i.e. the more cogent one. This condition holds even if there are absorptive media (as it is as is directly demonstrated in [40]). This can be heuristically understood because absorption means damping of the field values, which is a further favorable circumstance to control the growth of the unwanted solution fluctuations due to finite precision.

Conclusions

The work here presented concerned the study and implementation of an “in-house” three-dimensional FDTD (Finite-Difference Time-Domain) parallelized code, having good robustness and scalability. It is based on the direct numerical solution of the Maxwell’s equations in the time-domain, for the classical electromagnetic analysis of complex plasmonic and photonic structures interacting with light. A frequency analysis is performed by means of a discrete Fourier transform made inline on the time-domain data. Its reliability has been tested against the known analytic solution for the scattering off a sphere having a complex electric permittivity. The present work starts on the ground of a previous one for the MSc Degree Thesis in Physics. New accomplishments of this work are: the implementation of a new class of boundary conditions for the FDTD truncation of the computational domain; the implementation of a method to describe temporally dispersive media (the theoretical basis of both these arguments are found in literature); the implementation of a “plane wave injector” with arbitrary propagation direction allowing also, by means of superposition and a suitable diffraction integral, the numerical representation of focused gaussian beams for a more realistic description of the impinging light; the parallelization of the FDTD code using the Message Passing Interface (MPI) which, by suitably proposed used defined MPI data structures give the code good scalability allowing it to run efficiently on large High Performance Computing (HPC) machines with hundreds of thousands of calculating cores. To demonstrate its effectiveness and versatility, the developed FDTD code has been applied to the scattering and absorption analysis of nanoparticles made of different metals and of different shapes and sizes, either isolated or arrayed, emebded in a passivating silica layer and in front of a silicon substrate. This study as potential photovoltaic applications for the possible improvement of their efficiency in collecting and converting the solar light. The same FDTD code has also been applied to the study of an opal photonic crystal, being able to reproduce the experimental spectral reflectance/transmittance properties, in particular the band gaps. The analysis of photonic crystals finds applications in low threshold laser,

optical waveguides and sensor design. For the latter, the optical properties are connected to the crystal structure and symmetry, and vary with its geometry. It is important thus to have availability of the electromagnetic field distribution inside the crystal at the various wavelengths as obtained with the FDTD code. This knowledge allows the proposal for a procedure by means of which to reconstruct the banded $\omega - k$ dispersion curve of the crystal, based on a spatial Fourier transform of those field data. In the present work it is also described a general method for reflectance/transmittance calculations, at an arbitrary angle of incidence and with both TE and TM polarizations, of a stacked system of an arbitrary number of dielectric (with complex permittivity) layers of given thickness. This method, based on a matrix expression of the transmitted and reflected amplitudes at the various interfaces, is used in a proposed homogenization procedure of the photonic crystal, in which its discontinuous structure is averaged to get an equivalent unidimensional multilayer system with identical optical properties.

Further developments of the material in the present work and which would be extremely useful when implemented are:

The analysis of the field distribution of a photonic crystal when illuminated at an arbitrary direction of incidence of the impinging beam of light. This requires a knowledge of the phase factor between opposite crystal faces when Periodic Boundary Conditions (PBCs) are applied, and would allow the excitation of modes of arbitrary wave vector \vec{k} inside it, thus permitting the generalization of the banded $\omega - k$ reconstruction procedure in the whole of the first Brillouin zone of the crystal, without limiting it to a particular direction. If the direction of incidence cannot be varied (or equivalently, but with more difficulty, rotating the crystal sample) the relevant Bloch modes inside the crystal are hard to excite with presumably large errors in the iper-surface reconstruction.

The extension of the homogenization procedure to lattices or macroparticle shapes with different symmetries than the cubic or spherical ones. These generalizations could be described with stacked dielectric layers having a tensorial electric permittivity, perhaps only along directions perpendicular to propagation direction.

The analysis of completely or partially disordered colloidal photonic structures (photonic glasses) by simulating a completely random distribution of macroparticles, each with a size roughly comparable with the incident wavelengths and with an eventual homogenized description.

Finally, a big challenge could be the coupling of the electromagnetic FDTD code with an atomic or molecular system (inside the crystal) quantum mechanically described through the numerical implementation of Time Dependent

Schrödinger Equation, giving raise to a semi-classical description of a dopant in laser gain media.

Bibliography

- [1] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems*, Elsevier/Academic Press, 2013, 2nd edition.
- [2] W. C. Chew, *Waves and Fields in Inhomogeneous Media*, IEEE Press, 1995.
- [3] M. Born and E. Wolf, *Principles of Optics - Electromagnetic theory of propagation, interference and diffraction of light*, Cambridge University Press, 7th (expanded) edition.
- [4] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge University Press, 2008.
- [5] J. A. Stratton, *Electromagnetic Theory*, Mc Graw-Hill, 1941.
- [6] J. D. Jackson, *Classical Electrodynamics*, Wiley, 1998, 3rd ed.
- [7] L. D. Landau and E. M. Lifšits, *Elettrodinamica dei Mezzi Continui*, Editori Riuniti/Edizioni MIR, 1986.
- [8] K. Cahill, *Physical Mathematics* (Cambridge University Press, 2013).
- [9] P. Yeh, *Optical Waves in Layered Media*, Wiley, 1988.
- [10] W. Gropp, E. L. Lusk, and A. Skjellum, *Using MPI. Portable Parallel Programming with the Message Passing Interface*, The MIT Press, 1999, 2nd ed.
- [11] W. Gropp, E. L. Lusk, and R. Thakur, *Using MPI-2. Advanced Features in the Message Passing Interface*, The MIT Press, 1999.
- [12] P. Pacheco, *Parallel Programming with MPI*, Morgan Kaufmann, San Francisco, 1997.

- [13] R. Courant, K. O. Friedrichs, and H. Lewy, “On the partial difference equations of mathematical physics,” *IBM Jour. of Res. and Dev.* **11**, 215–234 (1967). English translation of the original 1928 paper.
- [14] K. S. Yee, “Numerical solution of initial boundary value problems involving Maxwell’s equations in isotropic media,” *IEEE Trans. Antennas Propagat.*, vol. AP-14, 302–307 (1966).
- [15] A. Taflov and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method* Artech House, Boston, 2005, 3rd ed.
- [16] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic crystals. Molding the flow of light*, Princeton University Press, Princeton and Oxford, 2008, 2nd ed., Appendix B.
- [17] W. Nagourney, *Quantum Electronics for Atomic Physics*, Oxford University Press, 2010.
- [18] L. Novotny and B. Hecht, *Principle of Nano-Optics*, Cambridge University Press, 2008.
- [19] A. G. Kulikovskii., N. V. Pogorelov, and A. YU. Semenov, *Mathematical aspects of numerical solution of hyperbolic systems*, Chapman & Hall/CRC, 2001.
- [20] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*, Wiley-Interscience, 1967, 2nd ed.
- [21] A. Bayliss and E. Turkel, “Radiation boundary conditions for wave-like equations,” *Commun. Pure Appl. Math.*, vol. XXXIII, 707–725 (1980).
- [22] B. Engquist and A. Majda, “Absorbing boundary conditions for the numerical simulations of waves,” *Math. Comput.* vol. 31, 629–651 (1977).
- [23] G. Mur, “Absorbing Boundary Conditions for the Finite-Difference Approximation of the Time-Domain Electromagnetic-Field Equations,” *IEEE Trans. Electromagn. Compat.*, vol. EMC-23, 377–382 (1981)
- [24] R. L. Higdon, “Absorbing Boundary Conditions for Difference Approximations to the Multi-Dimensional Wave Equation,” *Math. Comput.*, vol. 47, 437–459 (1986).
- [25] R. L. Higdon, “Numerical Absorbing Boundary Conditions for the Wave Equation,” *Math. Comput.*, vol. 49, 65–90 (1987).

- [26] L. N. Trefethen, “Instability of difference models for hyperbolic initial boundary value problems,” *Commun. Pure Appl. Math.*, vol. XXXVII, 329–367 (1984).
- [27] L. N. Trefethen and L. Halpern, “Well-Posedness of One-Way Wave Equations and Absorbing Boundary Conditions,” *Math. Comput.*, vol. 47, 421–435 (1986).
- [28] R. Pontalti, L. Cristoforetti, and L. Cescatti, “The frequency dependent FD-TD method for multi-frequency results in microwave hyperthermia treatment simulation,” *Phys. Med. Biol.*, vol. 38, 1283–1298 (1993).
- [29] A. Vaccari, R. Pontalti, C. Malacarne, and L. Cristoforetti, “A robust and efficient subgridding algorithm for finite-difference time-domain simulations of Maxwell’s equations,” *J. Comput. Phys.*, vol. 194, 117–139 (2004).
- [30] R. Pontalti, J. Nadobny, P. Wust, A. Vaccari, and D. Sullivan, “Investigation of Static and Quasi-Static Fields Inherent to the Pulsed FDTD Method,” *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-50(8), 2022–2025 (2002).
- [31] A. Vaccari, A. Calà Lesina, L. Cristoforetti, and R. Pontalti “Parallel implementation of a 3D subgridding FDTD algorithm for large simulations,” *Progress in Electromagnetics Research*, vol. 120, 263–292 (2011).
- [32] Z. Sacks, D. M. Kingsland, R. Lee, and J. Lee, “A Perfectly Matched Anisotropic Absorber for Use as an Absorbing Boundary Condition,” *IEEE Trans. Antennas Propagat.*, vol. AP-43(12), 1460–1463 (1995).
- [33] W. C. Chew and W. H. Weedon, “A 3D Perfectly Matched Medium from Modified Maxwell’s Equations with Stretched Coordinates,” *Microw. Opt. Tech. Lett.*, vol. 7(13), 599–604 (1994).
- [34] J. A. Roden and S. D. Gedney, “Convolution PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media,” *Microwave and Optical Technology Letters* **27**(5), 334–339 (2000).
- [35] D. Vasilevska, S. M. Goodnick, and G. Klimeck, *Computational Electronics. Semiclassical and Quantum Device Modeling and Simulation*, CRC Press, 2010.
- [36] M. Kuzuoglu and R. Mittra, “Frequency Dependence of the Constitutive Parameters of Causal Perfectly Matched Anisotropic Absorbers,” *IEEE Microwave and Guided Wave Letters* **6**(12), 447–449 (1996).

- [37] R. J. Luebbers, F. P. Hunsberger, K. S. Kunz, R. B. Standler, and M. Schneider, "A Frequency-Dependent Finite-Difference Time-Domain Formulation for Dispersive Materials," *IEEE Trans. Electromagnetic Compat.*, vol. EMC-32(3), 222–227 (1990).
- [38] R. J. Luebbers and F. P. Hunsberger, "FDTD for Nth-Order Dispersive Media," *IEEE Trans. Antennas Propagat.*, vol. AP-40(11), 1297–1301 (1992).
- [39] I. Laasko, S. Ilvonen, and T. Uusitupa, "Performance of convolutional PML absorbing boundary conditions in finite-difference time-domain SAR calculations," *Phys. Med. Biol.*, vol. 52, 7183–7192 (2007).
- [40] A. Vaccari, "Calcolo del campo elettromagnetico ad alta frequenza in strutture complesse. Applicazione radioprotezionistica per il telefono cellulare," MSc degree thesis in Physics, University of Trento, 1995.
- [41] İ. R. Çapoğlu, A. Taflove, and V. Backman, "Generation of an incident focused light pulse in FDTD," *Opt. Express*, vol. 16(23), 19208–19220 (2008).
- [42] İ. R. Çapoğlu, A. Taflove, and V. Backman, "Computation of tightly-focused laser beams in the FDTD method," *Opt. Express*, vol. 21(1), 87–101 (2013).
- [43] E. Wolf, "Electromagnetic Diffraction in Optical Systems. I. An Integral Representation of the Image Field," *Proc. R. Soc. Lond. A*, 253, 349–357 (1959).
- [44] B. Richards and E. Wolf, "Electromagnetic Diffraction in Optical Systems. II. Structure of the Image Field in an Aplanatic System," *Proc. R. Soc. Lond. A*, 253, 358–379 (1959).
- [45] A. Deinega and S. John, "Effective optical response of silicon to sunlight in the finite-difference time-domain method," *Opt. Letters*, vol. 37(1), 112–114 (2012).
- [46] K. P. Prokopidis and D. C. Zografopoulos, "A Unified FDTD/PML Scheme Based on Critical Points for Accurate Studies of Plasmonic Structures," *J. Lightwave Tech.*, vol. 31(15), 2467–2476 (2013).
- [47] K. P. Prokopidis and D. C. Zografopoulos, "Investigation of the Stability of ADE-FDTD methods for Modified Lorentz Media," *IEEE Microw. Wirel. Comp. Lett.*, vol. 24(10), 659–661 (2014).

- [48] A. Vial and T. Laroche, “Comparison of gold and silver dispersion laws suitable for FDTD simulations,” *App. Phys. B*, vol. 93, 139–143, (2008).
- [49] A. Paris, A. Vaccari, A. Calà Lesina, E. Serra, and L. Calliari “Plasmonic Scattering by Metal Nanoparticles for Solar Cells,” *Plasmonics* **7**, 525–534 (2010).
- [50] D. Sarid and W. Challener, *Modern Introduction to Surface Plasmons. Theory, Mathematica Modeling, and Applications*, Cambridge University Press, 2010.
- [51] H. A. Atwater and A. Polman “Plasmonics for improved photovoltaic devices,” *Nat. Mater.* **9**, 205–213 (2010).
- [52] C. F. Bohren and D. R. Huffman, *Absorption and scattering of light by small particles*, Wiley, 1998.
- [53] E. D. Palik, *Handbook of optical constants of solids*, Academic Press, 1985.
- [54] On the Web at: <http://refractiveindex.info>
- [55] J. R. Cole and N. J. Halas “Optimized plasmonic nanoparticle distributions for solar spectrum harvesting,” *Appl. Phys. Lett.* **89**, 153120 (2006).
- [56] *Reference Solar Spectral Irradiance: Air Mass 1.5*. On the Web at: <http://rredc.nrel.gov/solar/spectra/am1.5/>
- [57] A. Chiappini, C. Armellini, A. Chiasera, M. Ferrari, Y. Jestin, M. Mattarelli, M. Montagna, E. Moser, G. Nunzi Conti, G. C. Righini, M. Clara Gonçalves, and R. M. Almeida “Design of photonic structures by sol-gel-derived silica nanospheres,” *J. Non-Cryst. Solids* **353**, 674–678 (2007).
- [58] A. Chiappini, C. Armellini, A. Chiasera, M. Ferrari, L. Fortes, M. C. Gonçalves, R. Guider, Y. Jestin, R. Retoux, G. N. Conti, S. Pelli, R. M. Almeida, and G. C. Righini, “An alternative method to obtain direct opal photonic crystal structures,” *J. Non-Cryst. Solids* **355**, 1167–1170 (2009).
- [59] A. Vaccari, L. Cristoforetti, A. Calà Lesina, L. Ramunno, A. Chiappini, F. Prudenzeno, A. Bozzoli, and L. Calliari, “Parallel FDTD modeling of an opal photonic crystal,” *Optical Engineering* **53**(7), 071809 (2014).

- [60] A. Vaccari, A. Calà Lesina, L. Cristoforetti, A. Chiappini, L. Crema, L. Calliari, L. Ramunno, P. Berini, and M. Ferrari, “Light-opals interaction modeling by direct numerical solution of Maxwell’s equations,” *Optics Express* **22**(22), 27739–27749 (2014).