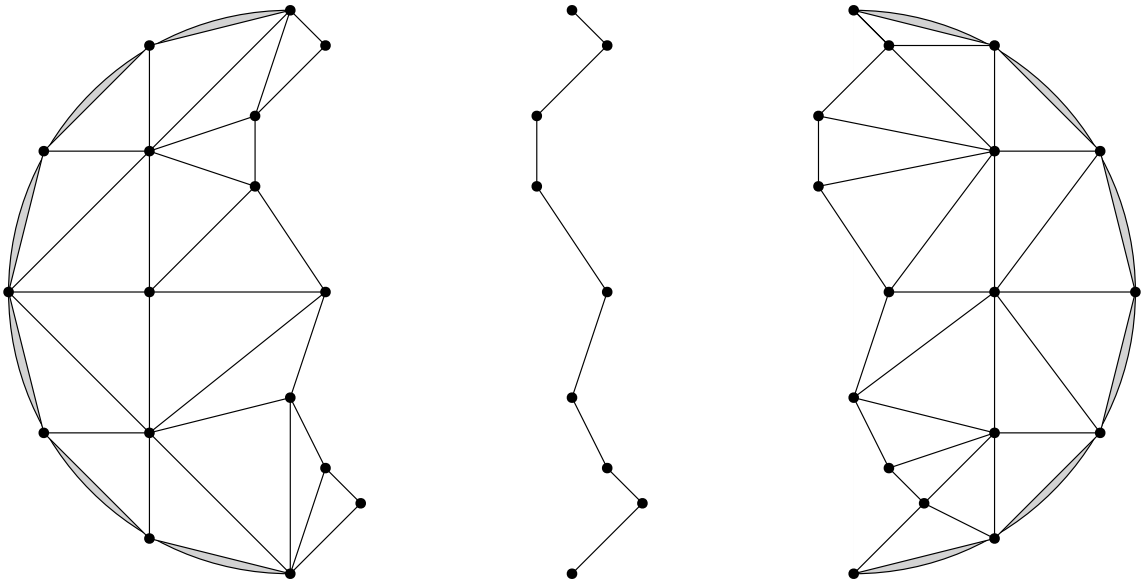
University of Liège
Faculty of Applied Sciences

Group of Electrical Engineering, Paris

Phased array optimization based on non-overlapping domain decomposition method using finite elements

Graduation Studies conducted to obtain the
Master's degree in Physical Engineering

Academic year 2017-2018



Author: Xavier ADRIAENS

Supervisor: Lionel PICHON — Directeur de recherche (GeePs)
Advisors: Abelin KAMENI NTICHI — Maître de conférences (GeePs)
Bertrand THIERRY — Chercheur permanent (LJLL)

Acknowledgement

I would like to thank Bertrand Thierry for the numerous individual lessons on the theory and numerical implementation of domain decomposition methods he taught me, starting from scratch. I also thank Abelin Kameni Nitchi for his everyday help, from sharing his computational and theoretical electromagnetics knowledge to debugging my programming mistakes through productive discussions on the research directions to take.
Of course I thank Lionel Pichon for welcoming me warmly to his research group but also for guiding this work and providing connections with other researchers.

Then I thank Antony Royer and Christophe Geuzaine for rapidly updating the software that are used in this work when necessary.

Finally I thank Julian Dular and Benjamin Benjamin De Bruyne for their careful proofreading and their pertinent comments.

Contents

Introduction 2

I Domain decomposition methods applied to the Helmholtz equation and their finite element discretization 3

1 Helmholtz equation in electromagnetism 4
- 1.1 Maxwell's equations . . . 4
- 1.2 Radiation conditions . . . 6
- 1.3 Total, incident and scattered fields . . . 7
- 1.4 Impenetrable scatterer . . . 9
- 1.5 Mono-dimensional example . . . 10
- 1.6 Domain truncation . . . 11
  - 1.6.1 Exact absorbing boundary condition . . . 11
  - 1.6.2 Zeroth order absorbing boundary condition . . . 13
  - 1.6.3 Second order absorbing boundary condition . . . 14

2 Non-overlapping Optimized Schwarz algorithms 16
- 2.1 Domain decomposition . . . 16
- 2.2 Equivalent formulation as a surface problem . . . 17
- 2.3 Iterative schemes . . . 19
  - 2.3.1 Jacobi method . . . 20
  - 2.3.2 Krylov methods . . . 21
- 2.4 Transmission condition . . . 21
  - 2.4.1 Classical Schwarz operator . . . 23
  - 2.4.2 Dirichlet-to-Neumann based operators . . . 24
    - (a) Real wavenumber . . . 25
    - (b) Complex wavenumber . . . 27

3 Nodal finite element discretization 31
- 3.1 Weak forms . . . 31
- 3.2 Discretization . . . 32

II Phased array optimization in heterogeneous media 36

4 Phased arrays 37
- 4.1 Definition . . . 37
- 4.2 Point source arrays . . . 38

4.2.1 Far field in free space . . . 39
4.3 Point source as a physical limit . . . 42

5 Optimization problems **45**
5.1 Problem based on power conservation . . . 45
5.2 Problems based on equivalent arrays . . . 48
5.3 Solution techniques . . . 49
5.3.1 Pure quadratic equality constrained problems . . . 49
5.3.2 Quadratic unconstrained problems . . . 49

III Applications **50**

6 Numerical validation of elementary solution **51**
6.1 Absorbing boundary conditions . . . 51
6.2 Point source approximation . . . 55
6.3 Transmission condition and iterative solver . . . 57
6.3.1 Real wavenumber . . . 57
6.3.2 Complex wavenumber . . . 58

7 Numerical validation of optimization **62**
7.1 Total outgoing power maximization . . . 62
7.2 Directive outgoing power maximization . . . 66

8 Power transfer through biological tissues **68**
8.1 Biological tissues model . . . 68
8.2 Specific absorption rate based constraints . . . 71
8.3 Phased array configuration . . . 72
8.4 Optimization . . . 73

Conclusion **76**

A Conventions and notations **80**

B Additional proofs, developments and theorems **83**
B.1 Dirichlet to Neumann (DtN) operator for a model problem . . . 83
B.2 Reflection coefficient for curved boundaries . . . 84
B.3 Divergence theorem . . . 86
B.4 Weak formulations . . . 86
B.4.1 Volume problem . . . 86
B.4.2 Surface equation . . . 88
B.4.3 Boundary operators . . . 88
B.5 Mass and stiffness matrices definitions . . . 88
B.6 Link between time and frequency domains for time harmonic fields . . . 89
B.7 Poynting's theorem . . . 90
B.7.1 Time domain . . . 90
B.7.2 Frequency domain . . . 91

B.8 Duality between losses minimization and radiation maximization 93
B.9 Link between electric losses and thermal elevation 93

C Additional tables and figures **95**

Introduction

Antennas design is a major concern when it comes to wireless electromagnetic systems. Power transfer, mobile communications or object identification (*e.g.* RADAR and LiDAR) are good examples among lots of others. For all these applications, it is critical to send the information with enough energy and only to the location of interest so that energy waste, electromagnetic pollution and human exposure are limited.
As any electromagnetic system, antennas behaviour is dictated by the Maxwell's equations. The behaviour of antennas at a particular frequency being an important parameter, the frequency domain Maxwell's equations are considered when possible. Exact solutions to these equations exist only when the media and the antenna geometry are simple, *e.g.* for stratified homogeneous media, for perfectly conducting spheres *etc*. When the complexity of the media or the antenna is high, numerical simulations become mandatory. However, solving numerically these equations is particularly difficult in this case as no terms in these equations are negligible, *i.e.* the *full wave* solution is needed as wave propagation is studied. The numerical challenge is that the Maxwell (or Helmholtz) differential operator is indefinite which implies that iterative solvers for volume finite elements (or finite differences) do not converge [9]. Expensive direct solver must then be used. The cost of these methods, typically $\mathcal{O}(N^3)$ where N is the number of unknowns, makes them impossible to apply practically for electrically large problem.
Several solutions to this numerical problem have therefore been provided [9]. One of them being the so called domain decomposition method. As its name indicates, this method consists in dividing the resolution volume into several (N_d) subvolumes with corresponding subproblems. Then the expensive exact solver is applied for the subproblems of each subvolumes. The challenge is of course to define subproblems that are equivalent to the overall original problem which can not be done in one step thus leading to an iterative technique. Using this method, the complexity goes from $\mathcal{O}(N^3)$ to $\mathcal{O}(N_d n_{\text{it}}(N/N_d)^3)$ but the main advantage is that the subproblems can be made independent for each iterations and thereforee solved by different computers. Aside that a parallel solution is faster, it also prevents memory problem that could occur when handling all the unknowns on the same computer. Applying domain decomposition method to antenna design could therefore open new possibilities as the numerical simulations of complex configurations become possible.

The subject of this work is to apply domain decomposition method to antenna design. It is divided into three parts.
The first part is dedicated to the *non-overlapping Schwarz algorithm*, a particular type of domain decomposition method and its application to typical antenna problem formulations. For the simplicity of the derivation, the particular case of transverse fields is considered. A word is also given on the finite element implementation of this algorithm.
The second part introduces the antenna geometries that are considered, *i.e.* phased arrays, and the optimization problems related to these geometries. Phased arrays have been chosen because they are

the simplest antennas, because their behaviour is quite intuitive and because the derivation of this part can be easily generalized to real antennas.
The third and last part applies the results from the two previous ones. First, the error made by the different approximations on the Maxwell's equations, and in particular the domain decomposition algorithm, are quantified on trivial cases. Then the antenna optimization is performed on two cases representing typical problem of power transfer through human tissue.

The algorithms described in this work are implemented using the finite element framework provided by gmsh[13], getDP[8] and getDDM[28] while the post-processing is done using Julia[2].
The codes related to this work are available online on the git repository `https://github.com/XavierAdriaens/Phased_arrays.git`.

Part I

Domain decomposition methods applied to the Helmholtz equation and their finite element discretization

Chapter 1

Helmholtz equation in electromagnetism

This chapter introduces the equation to which domain decomposition method is applied, *i.e.* the scalar Helmholtz equation, which is a particular case of the more general Maxwell's equations.

1.1 Maxwell's equations

Consider the *time domain* Maxwell's equations

$$\begin{aligned}
\textbf{curl}\ \mathbf{h} &= \mathbf{j} + \frac{\partial \mathbf{d}}{\partial t} && \text{(Maxwell-Ampère)}, && (1.1)\\
\textbf{curl}\ \mathbf{e} &= -\frac{\partial \mathbf{b}}{\partial t} && \text{(Faraday)}, && (1.2)\\
\text{div}\ \mathbf{b} &= 0 && \text{(Gauss)}, && (1.3)\\
\text{div}\ \mathbf{d} &= \rho && \text{(Coulomb)} && (1.4)
\end{aligned}$$

where $\mathbf{h}$ [A/m] is the *magnetic field*, $\mathbf{e}$ [V/m] is the *electric field*, $\mathbf{b}$ [W/m^2] is the *magnetic flux density*, $\mathbf{d}$ [C/m^2] is the *electric displacement field*, $\mathbf{j}$ [A/m^2] is the *current density* and ρ [C/m^3] is the *charge density*.
Substituting all these quantities by their Fourier transform[1]

$$\mathbf{f} = \int_{-\infty}^{\infty} \boldsymbol{f} \exp(i\omega t)\, d\omega \tag{1.6}$$

with ω [rad/s] the *pulsation*, yields the *frequency domain* Maxwell's equations

$$\begin{aligned}
\textbf{curl}\ \boldsymbol{h} &= \boldsymbol{j} + i\omega \boldsymbol{d} && \text{(Maxwell-Ampère)}, && (1.7)\\
\textbf{curl}\ \boldsymbol{e} &= -i\omega \boldsymbol{b} && \text{(Faraday)}, && (1.8)\\
\text{div}\ \boldsymbol{b} &= 0 && \text{(Gauss)}, && (1.9)\\
\text{div}\ \boldsymbol{d} &= \rho && \text{(Coulomb)}. && (1.10)
\end{aligned}$$

[1] $\boldsymbol{f}$ has therefore units $[[\mathbf{f}]/(\text{rad/s})]$. However for *time harmonic* fields at pulsation ω_0, the frequency domain fields write as

$$\boldsymbol{f} = \frac{1}{2}\left(\boldsymbol{f}'(\omega_0)\delta(\omega-\omega_0) + \boldsymbol{f}'^{*}(\omega_0)\delta(\omega+\omega_0)\right) \tag{1.5}$$

and $\boldsymbol{f}'$ has units $[[\mathbf{f}]]$.
In this work $\boldsymbol{f}'$ and $\boldsymbol{f}$ (resp. ω and ω_0) are used indifferently. The only difference between both is their units.

To complete these equations, consider also the *linear* and *isotropic* constitutive laws but possibly *dispersive*[2]

$$\boldsymbol{j} = \boldsymbol{j}_c + \boldsymbol{j}_a \text{ with } \boldsymbol{j}_c = \sigma \boldsymbol{e}, \quad \boldsymbol{b} = \mu \boldsymbol{h} \quad \text{and} \quad \boldsymbol{d} = \epsilon \boldsymbol{e}, \tag{1.11}$$

where $\boldsymbol{j}_c$ and $\boldsymbol{j}_a$ are the *conduction* and the *applied* current densities, σ [S/m] is the *conductivity*, μ [H/m] is the *magnetic permeability* and ϵ [F/m] is the *electric permittivity*.

Helmholtz equation for the electric field In the particular case where the magnetic permeability is homogeneous (*i.e.* constant w.r.t. the position), denoting this permeability μ_0, the magnetic field can be expressed as

$$\boldsymbol{h} = \frac{-1}{i\omega\mu_0} \textbf{curl}\, \boldsymbol{e} \tag{1.12}$$

using Faraday's law (Eq.(1.8)). Then injecting this relation into Maxwell-Ampère's law (Eq.(1.7)), yields

$$\frac{-1}{i\omega\mu_0} \textbf{curl}\ (\textbf{curl}\, \boldsymbol{e}) = \boldsymbol{j}_a + \sigma \boldsymbol{e} + i\omega\epsilon \boldsymbol{e}. \tag{1.13}$$

Defining the *complex permittivity* as

$$\tilde{\epsilon} \triangleq \epsilon + \frac{\sigma}{i\omega} \tag{1.14}$$

and the *wavenumber* k [1/m^2] as

$$k^2 \triangleq \omega^2 \mu \tilde{\epsilon}, \tag{1.15}$$

Eq.(1.13) becomes

$$\boxed{\left(\textbf{curl curl}\ - k^2\right) \boldsymbol{e} = -i\omega\mu_0 \boldsymbol{j}_a} \tag{1.16}$$

which is called the *vectorial non-homogeneous Helmholtz equation* for the electric field.

It is common in electromagnetism that polarization losses appear in media. These losses can be modelled through an imaginary part of the permittivity. The *equivalent conductivity* is then defined as

$$\tilde{\sigma} \triangleq \sigma + \omega \text{Im}\left(\epsilon\right). \tag{1.17}$$

The imaginary part of the permittivity and thus of the wavenumber squared are then negligible if

$$\frac{\tilde{\sigma}}{\omega \text{Re}\left(\epsilon\right)} \ll 1. \tag{1.18}$$

These kinds of materials are referred to as *bad conductors* (BC). At the opposite, if

$$\frac{\tilde{\sigma}}{\omega \text{Re}\left(\epsilon\right)} \gg 1 \tag{1.19}$$

then the imaginary part of the permittivity and of the wavenumber squared are dominant. These kinds of materials are referred to as *good conductors* (GC). Finally, if none of the above relations are satisfied, then the material is referred to as *dielectric* (D).

[2] A dispersive quantity is a quantity that depends on the pulsation ω.

Helmholtz equation for the magnetic field Similarly, in the particular case where the equivalent conductivity and the complex permittivity are homogeneous, denoting them respectively by $\tilde{\sigma}_0$ and $\tilde{\epsilon}_0$, the electric field can be expressed as

$$\boldsymbol{e} = \frac{\textbf{curl}\, \boldsymbol{h} - \boldsymbol{j}_a}{i\omega\tilde{\epsilon}_0} \tag{1.20}$$

using Maxwell-Ampère's law (Eq.(1.7)). Then injecting this relation into the Faraday's law (Eq.(1.8)), it yields

$$\boxed{\left(\textbf{curl curl}\ - k^2\right)\boldsymbol{h} = \textbf{curl}\, \boldsymbol{j}_a} \tag{1.21}$$

which is called the *vectorial non-homogeneous Helmholtz equation* for the magnetic field.

Two dimensions scalar Helmholtz equation In this work, the concepts are explained and applied to a simpler equation than the two previously introduced. This simpler equation can be obtained from the vectorial ones, *i.e.* Eq.(1.16) and Eq.(1.21), by considering either *transverse electric* (TE) fields

$$\boldsymbol{e} = u(x,y)\hat{\boldsymbol{z}} \quad \text{and} \quad -i\omega\mu_0\boldsymbol{j}_a = f(x,y)\hat{\boldsymbol{z}} \tag{1.22}$$

or *transverse magnetic* (TM) fields

$$\boldsymbol{h} = u(x,y)\hat{\boldsymbol{z}} \quad \text{and} \quad \textbf{curl}\, \boldsymbol{j}_a = f(x,y)\hat{\boldsymbol{z}}. \tag{1.23}$$

Inserting these ansatz into respectively Eq.(1.16) and Eq.(1.21) yields

$$\boxed{\left(-\,\text{div}\,\textbf{grad}\ - k^2\right)u = f} \tag{1.24}$$

which is called the *scalar non-homogeneous Helmholtz equation*.

1.2 Radiation conditions

Both vectorial and scalar non-homogeneous Helmholtz equations introduced previously (*i.e.* Eq.(1.16), Eq.(1.21) and Eq.(1.24)) do not have a unique solution and a supplementary condition must be added to ensure the uniqueness of the solution. In [26] [30], the authors show[3] that the solution to Maxwell's equations (Eq.(1.8) to Eq.(1.10)) for homogeneous and isotropic medium ($\mu \to \mu_0$ and $\epsilon \to \epsilon_0$) must verify[4]

$$\lim_{\rho\to\infty} \rho\left(\hat{\boldsymbol{x}} \times \boldsymbol{h} + \sqrt{\frac{\epsilon_0}{\mu_0}}\boldsymbol{e}\right) = 0 \tag{1.25}$$

and

$$\lim_{\rho\to\infty} \rho\left(\hat{\boldsymbol{x}} \times \boldsymbol{e} - \sqrt{\frac{\epsilon_0}{\mu_0}}\boldsymbol{h}\right) = 0. \tag{1.26}$$

These two conditions are equivalent to

$$\boxed{\lim_{\rho\to\infty} \rho\left(\hat{\boldsymbol{x}} \times (\textbf{curl}\, \boldsymbol{e}) - ik_0\boldsymbol{e}\right) = 0} \tag{1.27}$$

[3] Actually in [26], the authors stated that not only Eq.(1.25) and Eq.(1.26) have to be verified but also that $\lim_{\rho\to\infty} \rho\boldsymbol{e}/\boldsymbol{h}$ must be finite. A condition that has later been showed to be a consequence of the vectorial Helmholtz equation Eq.(1.16) and Eq.(1.21) by [30].

[4] $\rho = \sqrt{x^2+y^2+z^2} = \|\boldsymbol{x}\|$

and

$$\boxed{\lim_{\rho\to\infty} \rho\left(\hat{\boldsymbol{r}} \times (\mathbf{curl}\, \boldsymbol{h}) - ik_0 \boldsymbol{h}\right) = 0} \tag{1.28}$$

using Faraday's law Eq.(1.12) and Maxwell-Ampère's law Eq.(1.20), with $\boldsymbol{j}_a = 0$ as it is assumed that applied currents vanish at infinity. These conditions are called *Silver-Müller radiation conditions*.
For the scalar case in D dimensions, the radiation condition is given by[5]

$$\boxed{\lim_{\rho\to\infty} \rho^{\frac{D-1}{2}} \left(\frac{\partial u}{\partial \rho} + ik_0 u\right) = 0} \tag{1.29}$$

which is called the *Sommerfeld radiation condition*.
The two dimentional radiation condition is very similar to what is obtained by introducing the ansatz Eq.(1.22) or Eq.(1.23) into the Silver-Müller radiation condition, *i.e.*

$$\lim_{\rho\to\infty} \rho \left(\frac{\partial u}{\partial \rho} + ik_0 u\right) = 0. \tag{1.30}$$

The only difference being that ρ is replaced by $\sqrt{\rho}$ as the dimension of the problem decreases from three to two. This difference comes from the integral representation theorem that is used to derive the radiation condition. This theorem typically involves the Green function associated to the Helmoltz equation. In two dimensions, this function decreases as $\rho^{-1/2}$ while as ρ^{-1} in three dimensions. In the following sections, only the scalar system

$$\boxed{\begin{cases} \left(-\operatorname{div} \mathbf{grad}\ - k^2\right) u = f \quad \forall \boldsymbol{x} \in \mathbb{R}^2, \\ \lim\limits_{\rho\to\infty} \sqrt{\rho} \left(\dfrac{\partial u}{\partial \rho} + ik_0 u\right) = 0. \end{cases}} \tag{1.31}$$

is considered explicitly but the concepts can be applied to vectorial systems Eq.(1.16) and Eq.(1.28) or Eq.(1.21) and Eq.(1.28). A treatment of the vectorial case can be found in [28].

1.3 Total, incident and scattered fields

Consider a particular case of Eqs(1.31) where the sources (*i.e.* regions of the space where f is non-vanishing) are localized in a closed subset Ω^s and where the wave number is homogeneous everywhere except in another closed subset Ω^d (the scatterer), as represented in Figure 1.1.

[5] $\rho = \sqrt{x^2 + y^2} = \|\boldsymbol{x}\|$

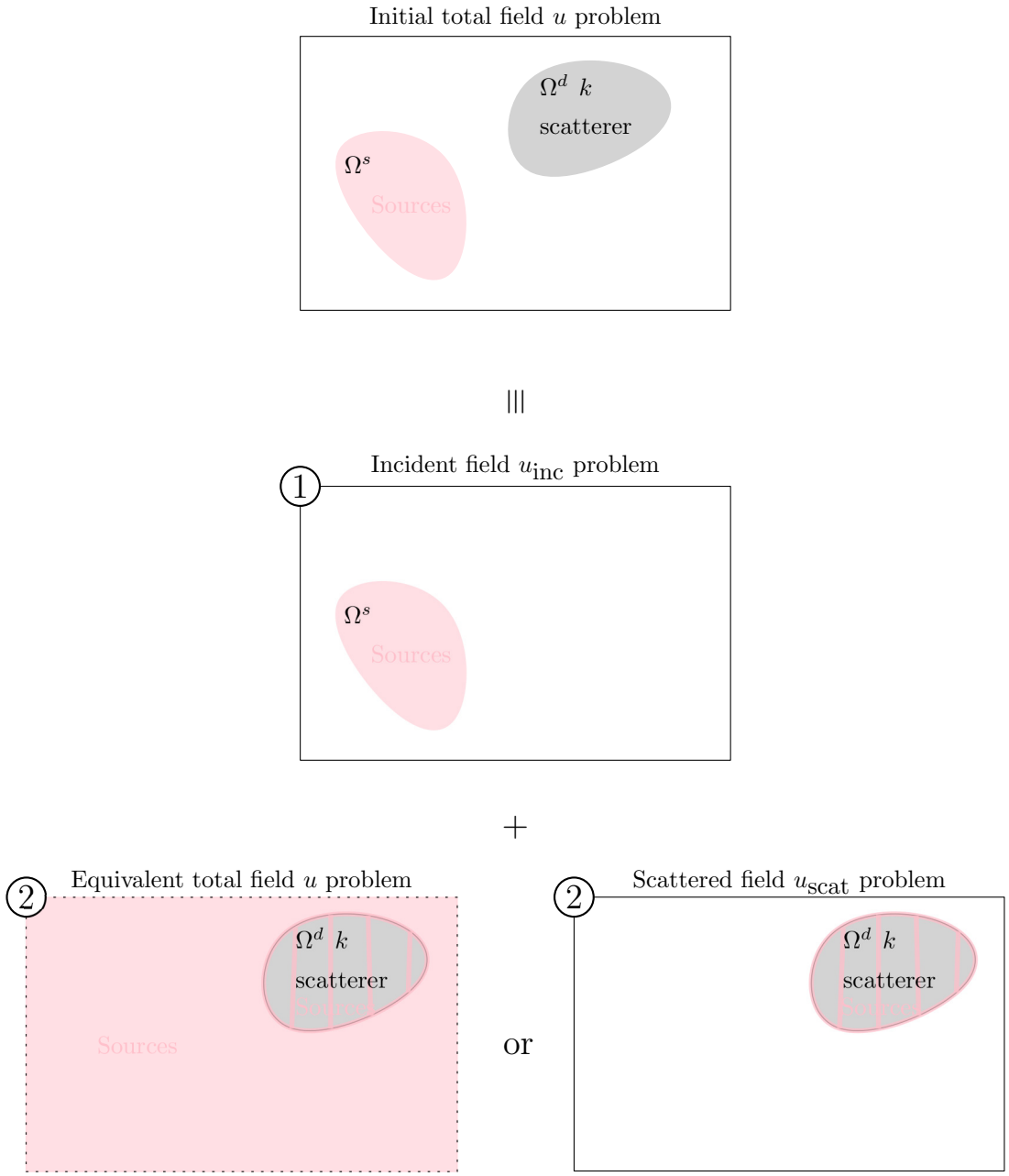


Figure 1.1: Illustration of the solution technique decomposing the total field u into an incident and scattered field u_{inc} and u_{scat}.

In such a case, the solution u can be divided as

$$u = u_{\text{inc}} + u_{\text{scat}} \tag{1.32}$$

where u_{inc} is the *incident* field and u_{scat} is the *scattered* field. The incident field is computed as the solution of

$$\begin{cases} \left(-\operatorname{div} \mathbf{grad} - k_0^2\right) u_{\text{inc}} = f \quad \forall \boldsymbol{x} \in \mathbb{R}^2, \\ \lim\limits_{\rho \to \infty} \sqrt{\rho} \left(\dfrac{\partial u_{\text{inc}}}{\partial \rho} + i k_0 u_{\text{inc}} \right) = 0. \end{cases} \tag{1.33}$$

The important point here is that the scatterer, *i.e.* the non-homogeneous region Ω^d, is not considered (*i.e.* $k = k_0$, $\forall \boldsymbol{x} \in \mathbb{R}^2$). The goal of this definition is to be able to formulate a new problem where

the sources f do not appear anymore and are subsituted by the field u_{inc} they generate. They are two possibilities to formulate this new problem, in terms of the total field u or in terms of the scattered field u_{scat}.

Scattered field formulation The equations verified by the scattered field are obtained by introducing the decomposition $u_{\text{scat}} = u - u_{\text{inc}}$ in the equations verified by the total field u, *i.e.* Eqs(1.31). Using then the definition of the incident field u_{inc} and the linearity of the Helmholtz operator, the scattered field equation is given by

$$\boxed{\begin{cases} \left(-\operatorname{div}\mathbf{grad} - k^2\right) u_{\text{scat}} = \left(k^2 - k_0^2\right) u_{\text{inc}} \quad \forall \boldsymbol{x} \in \mathbb{R}^2, \\ \lim\limits_{\rho\to\infty} \rho \left(\dfrac{\partial u_{\text{scat}}}{\partial \rho} + ik_0 u_{\text{scat}}\right) = 0. \end{cases}} \tag{1.34}$$

The important point is that the incident field u_{inc} acts as a fictitious source on the scatterer, where $k \neq k_0$.

Equivalent total field formulation Similarly, a formulation on the total field u where the sources f do not appear explicitly can be obtain by replacing f by $(-\operatorname{div}\mathbf{grad} - k_0^2)\, u_{\text{inc}}$ in Eqs(1.31). Sometimes however the incident field does not verify the Sommerfeld radiation condition, the most common example being the plane wave

$$u_{\text{inc}} = \exp\left(-i\boldsymbol{k}\cdot\boldsymbol{x}\right) \tag{1.35}$$

which is a very good approximation of the field radiated by any source at a point sufficiently far from it, *i.e.* in the so called *far field*. In that case, the total field can not satisfy the Sommerfeld radiation condition but only the diffracted field can. The system to solve for the total field is then

$$\boxed{\begin{cases} \left(-\operatorname{div}\mathbf{grad} - k^2\right) u = \left(-\operatorname{div}\mathbf{grad} - k_0^2\right) u_{\text{inc}} \quad \forall \boldsymbol{x} \in \mathbb{R}^2, \\ \lim\limits_{\rho\to\infty} \sqrt{\rho} \left(\dfrac{\partial u}{\partial \rho} + ik_0 u\right) = \lim\limits_{\rho\to\infty} \sqrt{\rho} \left(\dfrac{\partial u_{\text{inc}}}{\partial \rho} + ik_0 u_{\text{inc}}\right). \end{cases}} \tag{1.36}$$

One advantage of these two new formulations is that the considered domain can be restricted to the neighborhood of the scatterer. Moreover, sources often have the same topology (*e.g.* dipoles or windings), their radiating patterns (u_{inc}) are thus known and must not be computed to study every new scatterer.

1.4 Impenetrable scatterer

A common approximation in electromagnetism are *perfect conductors*. Inside a perfect conductor, the electric and magnetic fields $\boldsymbol{e}$ and $\boldsymbol{h}$ vanish. That is the reason why perfect conductors are also refered to as *impenetrable* scatterers. When the considered scatterer is a *perfect electric conductor* (PEC)(*i.e.* $\sigma \to \infty$), it can be shown that the electric and magnetic field satisfy

$$\boxed{\begin{cases} \hat{\boldsymbol{n}} \times \boldsymbol{e} = 0 \quad \forall \boldsymbol{x} \in \partial\Omega_{\text{PEC}}, \\ \hat{\boldsymbol{n}} \cdot \boldsymbol{h} = 0 \quad \forall \boldsymbol{x} \in \partial\Omega_{\text{PEC}}. \end{cases}} \tag{1.37}$$

Physically, these conditions impose that the tangential component of the electric field and the normal component of the magnetic field are continuous as both fields vanish inside the conductor.
By analogy with the perfect electric conductor, *perfect magnetic conductors* (PMC) are defined as a medium inside which fields vanish and such that

$$\boxed{\begin{cases} \hat{\boldsymbol{n}} \times \boldsymbol{h} = 0 & \forall \boldsymbol{x} \in \partial\Omega_{\text{PMC}}, \\ \hat{\boldsymbol{n}} \cdot \boldsymbol{e} = 0 & \forall \boldsymbol{x} \in \partial\Omega_{\text{PMC}}. \end{cases}} \tag{1.38}$$

An important point is that perfect magnetic conductors are only used for mathematical convenience[6] and not for modelling real common materials.
For the two dimensional scalar case, using the ansatz $\boldsymbol{e} = u(x,y)\hat{\boldsymbol{z}}$ or $\boldsymbol{h} = u(x,y)\hat{\boldsymbol{z}}$ and the modified Faraday's and Maxwell-Ampère's laws Eq.(1.12) and Eq.(1.20) into the PEC or PMC vectorial boundary conditions leads to the following boundary conditions

	$\boldsymbol{e} = u(x,y)\hat{\boldsymbol{z}}$	$\boldsymbol{h} = u(x,y)\hat{\boldsymbol{z}}$
$\forall \boldsymbol{x} \in \partial\Omega_{\text{PEC}}$	$u = 0,$ (1.39)	$\hat{\boldsymbol{n}} \cdot \mathbf{grad}\, u = 0,$ (1.40)
$\forall \boldsymbol{x} \in \partial\Omega_{\text{PMC}}$	$\hat{\boldsymbol{n}} \cdot \mathbf{grad}\, u = 0,$ (1.41)	$u = 0.$ (1.42)

1.5 Mono-dimensional example

To gain some intuition on the effect of the Sommerfeld radiation condition and on the scattered field formulation, consider the 1D problem of a plane wave $u_{\text{inc}} = \exp(-ik_0 x)$ incident on a perfectly conducting infinite plate stating at $x = 0$. The equations to consider in this case are given by Eqs(1.34) with the PEC boundary condition Eq.(1.39), *i.e.*

$$\begin{cases} \left(\partial_{xx}^2 - k_0^2\right) u_{\text{scat}} = 0 & \text{for } x < 0, \\ u_{\text{scat}} = -u_{\text{inc}} & \text{for } x \geq 0, \\ \lim\limits_{x \to -\infty} \left(-\dfrac{\partial u_{\text{scat}}}{\partial x} + ik_0 u_{\text{scat}} \right) = 0. \end{cases} \tag{1.43}$$

The solution to the first equation is

$$u_{\text{scat}} = A \exp(-ik_0 x) + B \exp(+ik_0 x) \tag{1.44}$$

i.e. the sum of an incoming and a reflected plane wave respectively. Physically, it is clear that the scattered field can only be composed of a reflected wave and thus that B should vanish. This is actually only the case if the Sommerfeld radiation condition is added, and this is precisely why this supplementary condition is needed, to prevent mathematical solutions that represents unphysical

[6] Except for some exotic materials.

situations where a scattered field carry energy from infinity towards the scatterer. Then using the Dirichlet boundary condition provides the expected solution

$$u_{\text{scat}} = \exp ik_0 x. \tag{1.45}$$

1.6 Domain truncation

The Helmholtz equation with Sommerfeld radiation condition Eqs(1.31) have to be solved[7] for all $\boldsymbol{x} \in \mathbb{R}^2$. In the scope of a numerical implementation that discretizes the resolution domain, *e.g.* finite element or finite difference methods, this resolution domain can not be unbounded as it would imply an infinite number of unknowns. Consequently, the resolution domain is restricted to a bounded open subset $\Omega \in \mathbb{R}^2$ whose boundary is denoted by $\partial\Omega$.
The radiation condition prescribes the asymptotic behavior of the solution at infinity. However, after domain truncation from $\mathbb{R}^2$ to Ω, infinity does not belong to the resolution domain anymore such that it is impossible to impose the radiation condition explicitly. For the problem to be well-posed, it is also necessary to impose a boundary condition on the new fictitious surface $\partial\Omega$. Ideally, this fictitious new boundary (and the related boundary condition) should not influence the solution. That is, the solution on $\mathbb{R}^2$ with exact radiation condition and the solution on Ω with the new boundary condition (to be determined) should be the same.
In this section, this exact boundary condition and two of its approximation are derived for a truncation along the y axis. This section is base on [23].

1.6.1 Exact absorbing boundary condition

Consider the scalar Helmholtz equation with Sommerfeld radiation condition Eqs(1.31) for an homogeneous medium with a real wavenumber k_0[8]. The unique solution to this system is denoted by u^*. The domain is truncated to $]-\infty, 0[\times\mathbb{R}$ (with the x-axis origin chosen such that the sources are in the half space $x < 0$) and the boundary condition $\mathcal{L}(u) = 0$ on $\{0\} \times \mathbb{R}$, $\mathcal{L}$ being a linear operator, is added. The considered system is thus

$$\begin{cases} \left(-\operatorname{div}\mathbf{grad} - k_0^2\right) u = f \quad \text{in }]-\infty, 0[\times\mathbb{R}, \\ \mathcal{L}(u) = 0 \quad \text{on } \{0\} \times \mathbb{R}, \\ \lim\limits_{\rho\to\infty, x<0} \sqrt{\rho}\left(\dfrac{\partial u}{\partial \rho} + ik_0 u\right) = 0. \end{cases} \tag{1.46}$$

The challenge here is to find a linear operator $\mathcal{L}$ such that the solution u of Eqs(1.46) is equal to the solution u^* of the original untruncated problem Eqs(1.31). Choosing the operator $\mathcal{L}$ such that $\mathcal{L}(u^*) = 0$ on $\{0\} \times \mathbb{R}$ is a choice that has the desired effect. Indeed, in that case u and u^* both verify Eqs(1.46). By uniqueness of the solution, one then has $u = u^*$. Taking $\mathcal{L}(u) = u - u^*(0, y)$ works but it is of course impossible to use practically as u^* is not known.
Another choice that works by construction is to take[9]

$$\boxed{\mathcal{L} = \hat{\boldsymbol{n}}_+ \cdot \mathbf{grad} + \Lambda} \tag{1.47}$$

[7] For all $\boldsymbol{x} \in \mathbb{R}^2 \setminus \overline{\Omega_{\text{PEC}}} \bigcap \overline{\Omega_{\text{PMC}}}$ if impenetrable scatterers are present.
[8] Most of the developments remain correct of the wavenumber as an imaginary part.
[9] $\hat{\boldsymbol{n}}_+ = \hat{\boldsymbol{x}}$

where Λ is the *Dirichlet-to-Neumann* (DtN) operator. This condition is called the *exact absorbing boundary condition*. The DtN operator is defined by[10]

$$\boxed{\begin{aligned} &\Lambda(g) = \hat{\boldsymbol{n}}_- \cdot \mathbf{grad}\, w \quad \text{on } \{0\} \times \mathbb{R} \\ &\text{with } w \text{ solution of} \\ &\begin{cases} \left(-\operatorname{div} \mathbf{grad}\, - k_0^2\right) w = 0 & \text{in }]0, \infty[\times \mathbb{R}, \\ w = g & \text{on } \{0\} \times \mathbb{R}, \\ \lim\limits_{\rho \to \infty, x > 0} \sqrt{\rho} \left(\dfrac{\partial w}{\partial \rho} + i k_0 w \right) = 0. \end{cases} \end{aligned}} \tag{1.48, 1.49}$$

Hence

$$\Lambda(u^*(0,y)) = \hat{\boldsymbol{n}}_- \cdot \mathbf{grad}\, w = \hat{\boldsymbol{n}}_- \cdot \mathbf{grad}\, u^* \tag{1.50}$$

by unicity as u^* and w both satisfy Eqs(1.49). One thus has

$$\mathcal{L}(u^*) = 0 \quad \text{on } \{0\} \times \mathbb{R} \tag{1.51}$$

with basic manipulations on Eq.(1.50).

As its name indicates, the DtN is an operator that takes the value of the field on the boundary (*i.e.* a Dirichlet data) and outputs the normal component of the gradient (*i.e.* a Neumann data). It is showed in Appendix B.1 that in this simple case, the DtN takes the form below

$$\boxed{\begin{aligned} &\Lambda(g) = \mathcal{F}^{-1}\left(\lambda \mathcal{F}(g)\right) \\ &\text{with } \lambda \text{ the } \textit{principal symbol} \text{ of } \Lambda \text{ given by} \\ &\lambda(k_y) = \begin{cases} \sqrt{k_y^2 - k_0^2} & \text{for } |k_y| \geq |k_0|\,, \\ j\sqrt{k_0^2 - k_y^2} & \text{for } |k_y| \leq |k_0| \end{cases} \end{aligned}} \tag{1.52, 1.53}$$

where $\mathcal{F}$ denotes the spatial Fourier transform w.r.t. y and k_y the spatial frequency.
The Λ operator is however *non local*, which means that the computation of $\Lambda(f)|_{y=y_0}$ depends on $f(y)\ \forall y \in \mathbb{R}$. This non locality actually comes from the principal symbol λ which involves a square root of the Fourier variable k_y[11].
Theoretically, the fact that Λ (and consequently $\mathcal{L}$) are non local is not a problem. It becomes problematic when considering a discretization of this operator because non local operators typically generate dense matrices. In the following subsections, two approximations of this exact absorbing boundary condition, yielding local operators, are presented. The idea to obtain a local operator is to express the principal symbol λ as a polynomial because

$$\mathrm{Op}(k_y^\alpha) = \mathcal{F}^{-1}\left(k_y^\alpha \mathcal{F}(\,\cdot\,)\right) = i^{-\alpha} \frac{\partial^\alpha}{\partial^\alpha y}, \quad \forall \alpha \geq 0 \tag{1.54}$$

is a local operator.

[10] $\hat{\boldsymbol{n}}_- = -\hat{\boldsymbol{x}}$

[11] It is useful here to remember that $\mathcal{F}^{-1}(k_y \hat{f}) = \dfrac{\partial f}{\partial y}$ is a local operator and at the opposite, that $\mathcal{F}^{-1}(k_y^{-1} \hat{f}) = \displaystyle\int_{-\infty}^{y} f(y) dy$ is typically a non local operator.

1.6.2 Zeroth order absorbing boundary condition

The zeroth order absorbing boundary condition is obtained by considering the zeroth order Taylor expansion of the principal symbol $\lambda(k_y)$ around $k_y \approx 0$ which writes

$$\begin{aligned} \lambda(k_y) &= i\sqrt{k_0^2 - k_y^2} \approx ik_0 \\ &\Rightarrow \Lambda \approx ik_0 \\ &\Rightarrow \mathcal{L}^0(u) = \hat{\boldsymbol{n}} \cdot \mathbf{grad}\, u + ik_0 u. \end{aligned} \tag{1.55}$$

It is interesting to point out that approximation Eq.(1.55) is equivalent to impose the Sommerfeld radiation condition Eq.(1.29) at a finite distance rather than at infinity.

Reflection coefficient To have some insights of the performance of the approximation Eq.(1.55), the effect of the absorbing boundary condition on the plane wave

$$u = \exp\left(-i(k_x x + k_y y)\right) \tag{1.56}$$

is considered. In the presence of an absorbing boundary condition along the y axis, a fraction of this wave is reflected such that the solution expresses as (*cfr* Section 1.5)

$$u = \exp\left(-i(k_x x + k_y y)\right) + \Gamma \exp\left(i(k_x x - k_y y)\right) \tag{1.57}$$

where Γ is determined by the boundary condition

$$\mathcal{L}^0(u) \triangleq \hat{\boldsymbol{n}} \cdot \mathbf{grad}\, u + ik_0 u = 0 \quad \text{on } x = 0, y \in \mathbb{R}. \tag{1.58}$$

After some manipulations, one obtains

$$\Gamma = \frac{k_x - k_0}{k_x + k_0} \tag{1.59}$$

which simplifies to

$$\Gamma = \frac{\cos\theta - 1}{\cos\theta + 1} \tag{1.60}$$

when introducing the incidence angle θ which is such that $k_x = k_0 \cos\theta$ and $k_y = k_0 \sin\theta$. The value of Γ is represented in Figure 1.2. As one can see, the reflection is small for $\theta \approx 0$, *i.e.* for normal incidence. It is important to highlight that it was expected that this absorbing boundary condition only works for small incidence as the approximation $k_y/k_0 \approx 0$ done in the Taylor expansion is equivalent to $\sin\theta \approx 0 \Leftarrow \theta \approx 0$.

Curved boundaries Eq.(1.55) can be generalized for curved boundaries. The zeroth order absorbing boundary condition then writes [19] as

$$\boxed{\mathcal{L}^0(u) = \hat{\boldsymbol{n}} \cdot \mathbf{grad}\, u + \left(ik_0 + \frac{\kappa}{2}\right) u} \tag{1.61}$$

where κ [1/m] is the local curvature. The reflection coefficient of plane waves on for curved boundaries is given in Appendix B.2.

Further in this work, the zeroth order boundary conditions are referred to as T0 for plane surfaces and CT0 for curved boundaries.

1.6.3 Second order absorbing boundary condition

Similarly, the second order absorbing boundary condition is obtained by considering the second order Taylor expansion of the principal symbol $\lambda(k_y)$ around $k_y \approx 0$, which writes

$$\begin{aligned}\lambda(k_y) = i\sqrt{k_0^2 - k_y^2} &\approx ik_0\left(1 - \frac{1}{2}\frac{k_y^2}{k_0^2}\right)\\ \Rightarrow \Lambda &\approx ik_0 - \frac{1}{2k_0 i}\frac{\partial^2}{\partial^2 y}\\ \Rightarrow \mathcal{L}^2(u) &= \hat{\boldsymbol{n}}\cdot\mathbf{grad}\, u + ik_0 u - \frac{1}{2k_0 i}\frac{\partial^2 u}{\partial^2 y}.\end{aligned} \tag{1.62}$$

Reflection coefficient As previously, the performance of the approximation Eq.(1.62) can be described by its effect on the plane wave Eq.(1.56). In the presence of an absorbing boundary condition along the y axis, a fraction of this wave is reflected such that the solution expresses again as Eq.(1.57) where Γ is determined by the boundary condition

$$\mathcal{L}^2(u) \triangleq \hat{\boldsymbol{n}}\cdot\mathbf{grad}\, u + ik_0 u - \frac{1}{2k_0 i}\frac{\partial^2 u}{\partial^2 y} = 0 \text{ on } x = 0, y \in \mathbb{R}. \tag{1.63}$$

After some manipulations, one obtains

$$\Gamma = \frac{k_x - k_0 + \frac{k_y^2}{2k_0}}{k_x + k_0 - \frac{k_y^2}{2k_0}} = \frac{\cos\theta - 1 + \frac{\sin^2\theta}{2}}{\cos\theta + 1 - \frac{\sin^2\theta}{2}} = -\left(\frac{\cos\theta - 1}{\cos\theta + 1}\right)^2. \tag{1.64}$$

The value of Γ is represented in Figure 1.2. As one can see, the reflection is small for $\theta \approx 0$, *i.e.* for normal incidence and as expected, smaller than for first order absorbing boundary condition.

Curved boundaries Eq.(1.62) can be generalized for curved boundaries. The second order absorbing boundary condition then writes [19] as

$$\boxed{\mathcal{L}^2(u) = \hat{\boldsymbol{n}}\cdot\mathbf{grad}\, u + \left(ik_0 + \frac{\kappa}{2} - \frac{\kappa^2}{8ik_0} - \frac{\kappa^3}{8k_0^2}\right)u - \left(\frac{1}{2ik_0} + \frac{\kappa}{2k_0^2}\right)\mathrm{div}_\Sigma\, \mathbf{grad}_\Sigma\, u.} \tag{1.65}$$

The reflection coefficient of plane waves on curved boundaries is given in Appendix B.2.

Further in this work, the second order boundary conditions are referred to as T2 for plane surfaces and CT2 for curved boundaries.
For the general Maxwell's equations, the equivalent for the DtN operator is the *Magnetic-to-Electric* (MtE) operator. Details about this operator can be found in [3].

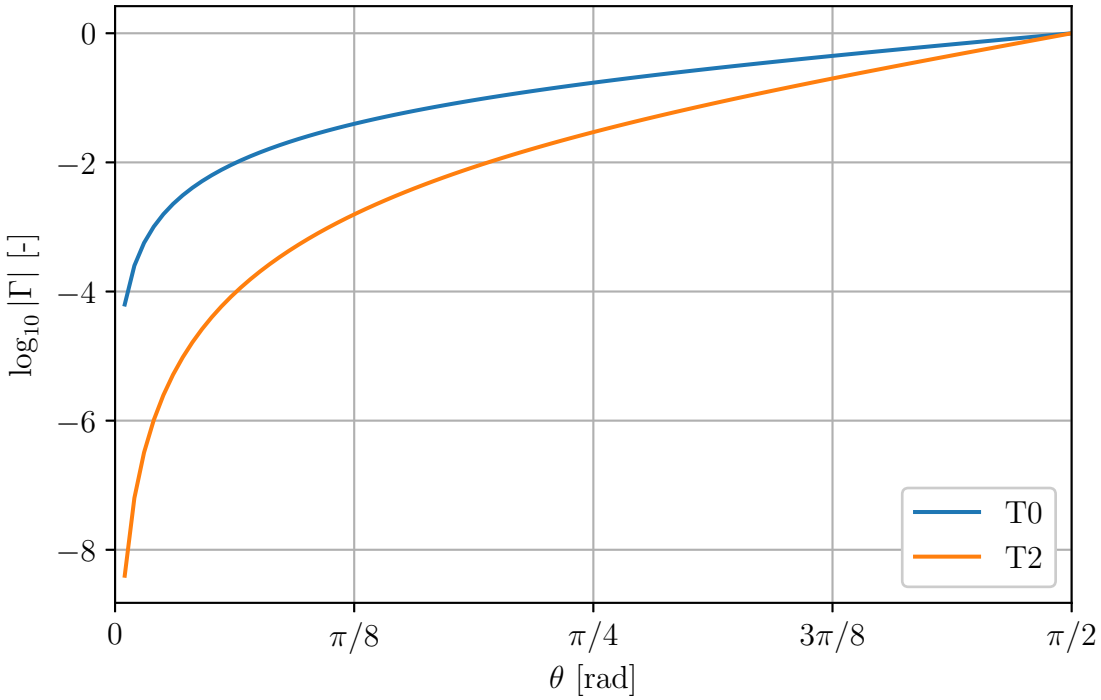


Figure 1.2: Reflection coefficient Γ for T0 Eq.(1.55) and T2 Eq.(1.62) for plane surfaces.

Chapter 2

Non-overlapping Optimized Schwarz algorithms

In the previous chapter, all the mathematical ingredients necessary to formulate typical electromagnetic waves problems were introduced. These problems can be summarized in the general form

$$\boxed{\begin{cases} \left(-\operatorname{div}\mathbf{grad} - k^2\right) u = f & \text{in } \Omega, \\ \left(\hat{\boldsymbol{n}}\cdot\mathbf{grad} + \mathcal{B}\right) u = u_\infty & \text{on } \partial\Omega^\infty, \\ u = u_d & \text{on } \partial\Omega^d, \\ \hat{\boldsymbol{n}}\cdot\mathbf{grad}\, u = u_n & \text{on } \partial\Omega^n \end{cases}} \tag{2.1}$$

where u can be both the scattered field or the total field, u_∞, u_d and u_n and f are known functions that can be deduced from results presented in the former chapter. The domain Ω is an open bounded subset of $\mathbb{R}^2$ whose boundary is denoted by $\partial\Omega$. This boundary is composed of three parts $\partial\Omega^\infty$, $\partial\Omega^d$ and $\partial\Omega^n$ which are respectively the boundary where absorbing, Dirichlet and Neumann boundary conditions are applied. $\mathcal{B}$ is a known (from Chapter 1) linear operator defined on $\partial\Omega^\infty$ that acts along its tangential direction only
The aim of this chapter is to present particular solution techniques for Eqs.(2.1) using specific kind of domain decomposition methods, *i.e.* non-overlapping Optimized Schwarz algorithms.

2.1 Domain decomposition

Instead of solving Eqs(2.1), non-overlapping domain decomposition methods suggest to solve the N_d (*i.e.* $\forall i \in [0, N_d - 1]$) coupled problems

$$\boxed{\begin{cases} \left(-\operatorname{div}\mathbf{grad} - k^2\right) u_i = f & \text{in } \Omega_i, \\ \left(\hat{\boldsymbol{n}}\cdot\mathbf{grad} + \mathcal{B}\right) u_i = u_\infty & \text{on } \partial\Omega_i^\infty, \\ u_i = u_d & \text{on } \partial\Omega_i^d, \\ \hat{\boldsymbol{n}}\cdot\mathbf{grad}\, u_i = u_n & \text{on } \partial\Omega_i^n, \\ \mathcal{J}_i u_i = \left(\mathcal{J}_i u_j \triangleq g_{ij}\right) & \text{on } \Gamma_{ij},\ \forall j \in D_i \end{cases}} \tag{2.2}$$

where $\{\Omega_i\, \forall i \in [0, N_d - 1]\}$ is a non-overlapping partition of Ω (*i.e.* $\Omega = \bigcup_i \Omega_i \bigcup_j \Gamma_{ij}$ but $\bigcap_i \Omega_i = \emptyset$) and where $\partial\Omega_i^\infty$, $\partial\Omega_i^n$ and $\partial\Omega_i^n$ are respectively a non-overlapping partition of $\partial\Omega^\infty$, $\partial\Omega^n$ and $\partial\Omega^n$

(*i.e.* $\partial\Omega_i^\infty = \partial\Omega^\infty \cap \partial\Omega_i$, $\partial\Omega_i^d = \partial\Omega^d \cap \partial\Omega_i$ and $\partial\Omega_i^n = \partial\Omega^n \cap \partial\Omega_i$). $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$ are the new interfaces created by the decomposition and $\mathcal{J}_i$ are the new boundary conditions linking the solutions of subdomains i and j on this new interface. D_i is the list containing the integer related to neighbouring subdomains of domain i, *i.e.* $D_i = \{k \in [0, N_d - 1] \,|\Gamma_{ik} \neq \emptyset\}$. These N_d problems are thus coupled through their boundaries.
In the following sections, f, u_∞, u_d, u_n will be referred to as the *physical sources* in opposition to the term $\mathcal{J}_i u_j \triangleq g_{ij}$ which represents a source term that appears only because of the domain decomposition but has no physical meaning and is thus sometimes referred to as the *artificial source*. The artificial source term g_{ij} is defined on the surface Γ_{ij} and is an unknown (because it depends on the volume unknown u_j) such that it is also called a *surface unknown*.

2.2 Equivalent formulation as a surface problem

None of the problems Eqs(2.2) can actually be solved independently as the artificial sources/surface unknowns $\mathcal{J}_i u_j \triangleq g_{ij}$ are unknown. In this section, a reformulation of the coupled equations Eqs(2.2) as a linear system is derived. The unknowns of this linear system are chosen to be the artificial sources/surface unknowns $g_{ij} \triangleq \mathcal{J}_i u_j$. This choice might be surprising because the quantity of interest is actually u_i, *i.e.* the volume unknowns. The reason for this choice will be highlighted later in this section. For now, just notice that knowing g_{ij}, Eqs(2.2) become uncoupled and can then be solved separately.
The first step towards this new system is the introduction of the operator $\mathcal{A}_{ji}$. The inputs of this operator are card(D_i) functions (s_{im}, one for all $m \in D_i$) each defined on a different artificial boundary Γ_{im} and its output is a single function $\mathcal{J}_j w$ defined on Γ_{ij}. The action of the operator is defined as

$$\mathcal{A}_{ji}(s_{im}\, \forall m \in D_i) = \mathcal{J}_j w, \tag{2.3}$$

with w solution of

$$\begin{cases} \left(-\text{div}\,\mathbf{grad}\, - k^2\right) w = f & \text{in } \Omega_i, \\ \left(\hat{\boldsymbol{n}} \cdot \mathbf{grad}\, + \mathcal{B}\right) w = u_\infty & \text{on } \partial\Omega_i^\infty, \\ w = u_d & \text{on } \partial\Omega_i^d, \\ \hat{\boldsymbol{n}} \cdot \mathbf{grad}\, w = u_n & \text{on } \partial\Omega_i^n, \\ \mathcal{J}_i w = s_{im} & \text{on } \Gamma_{im},\ \forall m \in D_i. \end{cases} \tag{2.4}$$

Physically this operator computes the volume solution w on subdomain i for a given set of artificial sources s_{im} (one source for each artificial boundary Γ_{im}) and outputs $\mathcal{J}_j w$. If the inputs s_{im} are the surface unknowns g_{im}, then $\mathcal{J}_j w$ is the artificial source coming from subdomain i to subdomain j, *i.e.* g_{ji}. By definition of $\mathcal{A}_{ji}$, Eqs.(2.2) can then be written as the following system

$$\left(\mathcal{J}_j u_i \triangleq g_{ji}\right) = \mathcal{A}_{ji}(g_{im}\, \forall m \in D_i),\ \forall i \in [0, N_d - 1]\,, \forall j \in D_i \tag{2.5}$$

whose unknowns are the artificial sources/surface unknowns g_{ij}.

Linearization The problem of the above system is that $\mathcal{A}_{ji}$ is not linear[1] w.r.t. the surface unknowns g_{ij} because of the physical source terms f, u_∞, u_d, u_n. However, the operator $\mathcal{A}$ can easily

[1]Consider for example $\mathcal{A}_{ji}\left(s_{im}^1 + s_{im}^2\, \forall m \in D_i\right)$ in the definition of $\mathcal{A}_{ji}$ Eq.(2.3) and Eqs.(2.4). It is straightforward to see that $\mathcal{A}_{ji}\left(s_{im}^1 + s_{im}^2\,, \forall m \in D_i\right) \neq \mathcal{A}_{ji}\left(s_{im}^1\,, \forall m \in D_i\right) + \mathcal{A}_{ji}\left(s_{im}^2\,, \forall m \in D_i\right)$ unless the physical sources f, u_∞, u_d, u_n vanish.

be decomposed into the sum of its linear part $\mathcal{A}'$ taking into account only the effect of the surface unknowns, and its constant (w.r.t. g_{ij}) part b taking into account only the physical sources.
Using $w = w_{\text{art}} + w_{\text{phy}}$ in Eq.(2.3) and Eqs.(2.4), $\mathcal{A}'_{ji}$ is given by

$$\boxed{\begin{array}{l}\mathcal{A}'_{ji}(s_{im}\,\forall m \in D_i) = \mathcal{J}_j w_{\text{art}}, \\ \text{with } w_{\text{art}} \text{ solution of}\end{array}} \tag{2.6}$$

$$\begin{cases} \left(-\operatorname{div}\mathbf{grad}\ - k^2\right) w_{\text{art}} = 0 & \text{in } \Omega_i, \\ \left(\hat{\boldsymbol{n}}\cdot\mathbf{grad}\ + \mathcal{B}\right) w_{\text{art}} = 0 & \text{on } \partial\Omega_i^{\infty}, \\ w_{\text{art}} = 0 & \text{on } \partial\Omega_i^{d}, \\ \hat{\boldsymbol{n}}\cdot\mathbf{grad}\ w_{\text{art}} = 0 & \text{on } \partial\Omega_i^{n}, \\ \mathcal{J}_i w_{\text{art}} = s_{im} & \text{on } \Gamma_{im},\ \forall m \in D_i \end{cases} \tag{2.7}$$

and b_{ji} by

$$\boxed{\begin{array}{l} b_{ji} = \mathcal{J}_j w_{\text{phy}}, \\ \text{with } w_{\text{phy}} \text{ solution of}\end{array}} \tag{2.8}$$

$$\begin{cases} \left(-\operatorname{div}\mathbf{grad}\ - k^2\right) w_{\text{phy}} = f & \text{in } \Omega_i, \\ \left(\hat{\boldsymbol{n}}\cdot\mathbf{grad}\ + \mathcal{B}\right) w_{\text{phy}} = u_{\infty} & \text{on } \partial\Omega_i^{\infty}, \\ w_{\text{phy}} = u_d & \text{on } \partial\Omega_i^{d}, \\ \hat{\boldsymbol{n}}\cdot\mathbf{grad}\ w_{\text{phy}} = u_n & \text{on } \partial\Omega_i^{n}, \\ \mathcal{J}_i w_{\text{phy}} = 0 & \text{on } \Gamma_{im},\ \forall m \in D_i. \end{cases} \tag{2.9}$$

This linear part can also be written compactly as

$$\mathcal{A}'_{ji}(s_{im}) = \mathcal{A}_{ji}(s_{im}; f = 0, u_{\infty} = 0, u_d = 0, u_n = 0) \tag{2.10}$$

and the constant part as

$$b_{ji} = \mathcal{A}_{ji}(s_{im} = 0; f, u_{\infty}, u_d, u_n). \tag{2.11}$$

The equation on the surface unknowns Eq.(2.5) then writes

$$g_{ji} = \mathcal{A}'_{ji}(g_{im}\,\forall m \in D_i) + b_{ji},\ \forall i \in [0, N_d - 1]\,, \forall j \in D_i. \tag{2.12}$$

To simplify the notations, the *surface unknowns vector*

$$\boxed{\boldsymbol{g} \triangleq \{g_{ij},\ \forall i \in [0, N_d - 1]\,, \forall j \in D_i\}} \tag{2.13}$$

is introduced such that the *linear system for the surface unknowns* is[2][3]

$$\boxed{\left(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{A}'}\right)\boldsymbol{g} = \boldsymbol{b}.} \tag{2.14}$$

[2] $\boldsymbol{b} \triangleq \{b_{ij},\ \forall i \in [0, N_d - 1]\,, \forall j \in D_i\}$
[3] $\boldsymbol{\mathcal{A}'} \triangleq \{\mathcal{A}'_{ij},\ \forall i \in [0, N_d - 1]\,, \forall j \in D_i\}$

Surface equation for a particular transmission condition The transmission operator (*cfr* section 2.4) $\mathcal{J}_i$ is often chosen to be of the form

$$\mathcal{J}_i = \hat{\boldsymbol{n}}_i \cdot \mathbf{grad} \; + \mathcal{S}_i \tag{2.15}$$

such that Eq.(2.3) becomes

$$\begin{aligned} \mathcal{A}_{ji}(s_{im}\,, \forall m \in D_i) &= (\hat{\boldsymbol{n}}_j \cdot \mathbf{grad} \; + \mathcal{S}_j)\, w \quad (\text{on } \Gamma_{ij}) \\ &= (-\hat{\boldsymbol{n}}_i \cdot \mathbf{grad} \; - \mathcal{S}_i + \mathcal{S}_i + \mathcal{S}_j)\, w \\ &= -\mathcal{J}_i w + (\mathcal{S}_i + \mathcal{S}_j)\, w \\ &= -s_{ij} + (\mathcal{S}_i + \mathcal{S}_j)\, w \end{aligned} \tag{2.16}$$

where the transmission boundary condition of Eqs(2.4) and the fact that $\hat{\boldsymbol{n}}_i = -\hat{\boldsymbol{n}}_j$ on Γ_{ij} have been used.

Difference with the volume unknowns system The previous reasoning could have been done without introducing the surface unknowns. Indeed, it is possible to write

$$u_i = \mathcal{A}_i^{\text{vol}}(u_m \, \forall m \in D_i) \tag{2.17}$$

with

$$\mathcal{A}_i^{\text{vol}}(s_m \, \forall m \in D_i) = w, \tag{2.18}$$

with w solution of

$$\begin{cases} \left(-\text{ div } \mathbf{grad} \; - k^2\right) w = f & \text{in } \Omega_i, \\ (\hat{\boldsymbol{n}} \cdot \mathbf{grad} \; + \mathcal{B})\, w = u_\infty & \text{on } \partial\Omega_i^\infty, \\ w = u_d & \text{on } \partial\Omega_i^d, \\ \hat{\boldsymbol{n}} \cdot \mathbf{grad}\, w = u_n & \text{on } \partial\Omega_i^n, \\ \mathcal{J}_i w = \mathcal{J}_i s_m & \text{on } \Gamma_{im},\ \forall m \in D_i. \end{cases} \tag{2.19}$$

As for the surface unknowns, the linear system for the volume unknowns would then write as

$$\left(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{A}'}^{\text{vol}}\right) \boldsymbol{u} = \boldsymbol{b}^{\text{vol}} \tag{2.20}$$

and is theoretically equivalent to Eq.(2.14). However when discretizing these equations, *e.g.* using finite element or finite difference methods, the size of the system to solve would be much greater in the case of volume unknowns as there are much more discrete unknowns related to a volume unknown field than an surface unknown field.

2.3 Iterative schemes

In this section, two iterative methods for solving the linear system Eq.(2.14) are used. The first is the Jacobi method. This method is presented because its iterative scheme has an intuitive interpretation and because convergence rates can be derived analytically for a two-subdomain model problem. The performance of this method are shown in Part III to be bad. That is why a second method is considered: the GMRES[24], a Krylov subspace method. Only the general principle of this method is given.

2.3.1 Jacobi method

The *iterative Jacobi* method applied to Eq.(2.14) consists in computing sequentially

$$\boxed{\boldsymbol{g}^{n+1} = \boldsymbol{\mathcal{A}'}\boldsymbol{g}^{n} + \boldsymbol{b}} \tag{2.21}$$

until the residual[4]

$$\boldsymbol{r}^{n} \triangleq \boldsymbol{b} - (\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{A}'})\,\boldsymbol{g}^{n} \tag{2.22}$$

is small enough, typically such that

$$\|\boldsymbol{r}^{n}\| \leq \epsilon_{\text{it}} \quad \text{or} \quad \frac{\|\boldsymbol{r}^{n}\|}{\|\boldsymbol{r}^{0}\|} \leq \epsilon_{\text{it}} \tag{2.23}$$

where ϵ_{it} is a user defined parameter and $\|\cdot\|$ is a norm suited to the problem, *e.g.* the $L_2(\Gamma_{ij})$ norm in this case.

Intuitive interpretation

From the definition of $\boldsymbol{\mathcal{A}}$, $\boldsymbol{\mathcal{A}'}$ and $\boldsymbol{b}$, a Jacobi iteration Eq.(2.21) is equivalent to

$$\boldsymbol{g}^{n+1} = \boldsymbol{\mathcal{A}}\boldsymbol{g}^{n}. \tag{2.24}$$

Then expliciting the $\boldsymbol{\mathcal{A}}$ operator gives

$$g_{ji}^{n+1} = \mathcal{A}_{ji}(g_{im}^{n}\,\forall m \in D_i) = \mathcal{J}_j w, \tag{2.25}$$

with w solution of

$$\begin{cases} \left(-\operatorname{div}\mathbf{grad} - k^2\right) w = f & \text{in } \Omega_i, \\ \left(\hat{\boldsymbol{n}}\cdot\mathbf{grad} + \mathcal{B}\right) w = u_\infty & \text{on } \partial\Omega_i^{\infty}, \\ w = u_d & \text{on } \partial\Omega_i^{d}, \\ \hat{\boldsymbol{n}}\cdot\mathbf{grad}\, w = u_n & \text{on } \partial\Omega_i^{n}, \\ \mathcal{J}_i w = g_{im}^{n} & \text{on } \Gamma_{im},\ \forall m \in D_i. \end{cases} \tag{2.26}$$

The intuitive interpretation of the preceding equation is that, at the $n+1$ iteration, first the solution to the volume problem Eqs(2.26) is computed taking into account the physical sources but also artificial sources coming from the neighbouring subdomains and resulting from the computation at the preceding iteration n. Then the surface unknown is updated through Eq.(2.25). The procedure is then repeated until convergence is reached.
It is important to highlight that Eq.(2.5) and Eq.(2.25) are very different because in the second case, there is no coupling between the problems to solve on each subdomain as the information from the preceding iteration is used.
An even more understandable way of writing Eq.(2.24), without any operator or surface unknowns

[4]In this case, the residual $\boldsymbol{r}^{n} = \boldsymbol{g}^{n+1} - \boldsymbol{g}^{n}$. As expected, when the residual vanish $\boldsymbol{g}^{n+1} = \boldsymbol{g}^{n}$.

[4]Numerically, this norm is substituted by the euclidean norm on the nodes unknowns because both norms are equivalent.

but only the volume unknowns, is

$$\begin{cases} \left(-\text{ div } \mathbf{grad} \ - k^2\right) u_i^{n+1} = f & \text{in } \Omega_i, \\ \left(\hat{\boldsymbol{n}} \cdot \mathbf{grad} \ + \mathcal{B}\right) u_i^{n+1} = u_\infty & \text{on } \partial\Omega_i^\infty, \\ u_i^{n+1} = u_d & \text{on } \partial\Omega_i^d, \\ \hat{\boldsymbol{n}} \cdot \mathbf{grad} \ u_i^{n+1} = u_n & \text{on } \partial\Omega_i^n, \\ \mathcal{J}_i u_i^{n+1} = \mathcal{J}_i u_m^n & \text{on } \Gamma_{im}, \ \forall m \in D_i. \end{cases} \tag{2.27}$$

This system is actually the simplest way of decoupling the original set of volume problems, Eqs(2.2).

2.3.2 Krylov methods

The idea of *Krylov subspace* methods is to search the solution of system Eq.(2.14) at the n^{th} iteration in a so called *Krylov subspace* $\mathcal{K}_n(\boldsymbol{\mathcal{A}'}, \boldsymbol{b})$. The solution $\boldsymbol{g}^n \in \mathcal{K}^n$ is then chosen such that it minimizes the residual $\boldsymbol{r}^n$ [24]. This Krylov subspace is defined as

$$\mathcal{K}_n(\boldsymbol{\mathcal{A}'}, \boldsymbol{b}) \triangleq \text{span}\{\boldsymbol{b}, \boldsymbol{\mathcal{A}'b}, \boldsymbol{\mathcal{A}'^2b}, \ldots, \boldsymbol{\mathcal{A}'^{n-2}b}, \boldsymbol{\mathcal{A}'^{n-1}b}\}. \tag{2.28}$$

This space is actually quite natural for searching the solution [18]. Indeed a corollary of Cayley-Hamilton theorem states that $\boldsymbol{\mathcal{A}'^{-1}}$ can be expanded as

$$\boldsymbol{\mathcal{A}'^{-1}} = \sum_{n=0}^{N-1} a_n \boldsymbol{\mathcal{A}'^n} \tag{2.29}$$

and therefore

$$\begin{aligned} &(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{A}'})\boldsymbol{g} = \boldsymbol{b} \\ \Rightarrow \quad &\boldsymbol{g} = (\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{A}'})^{-1} \boldsymbol{b} \\ \Rightarrow \quad &\boldsymbol{g} = \boldsymbol{b} - \sum_{n=0}^{N-1} a_n \boldsymbol{\mathcal{A}'^n b}. \end{aligned} \tag{2.30}$$

The GMRES algorithm is detailed in [24] while details about Krylov subspace solver can be found in [18].

2.4 Transmission condition

The *transmission conditions* are the supplementary boundary conditions that must be added when cutting the domain, because new boundaries Γ_{ij} appear. These conditions are one of the most important point in domain decomposition methods as they directly influence the convergence rate and hence the quality of the method. In the previous sections, these conditions were written in the general form

$$\mathcal{J}_i u_i = \left(\mathcal{J}_i u_j \triangleq g_{ij}\right) \text{ on } \Gamma_{ij}, \ \forall j \in D_i. \tag{2.31}$$

To be able to derive simple analytic conclusions about the transmission conditions, a model example is used. This section is based on [11].

Consider the Jacobi iterative scheme applied to the particular case of two semi-infinite homogeneous subdomains, without impenetrable scatterers, *i.e.* $\Omega_i^n = \Omega_i^d = \emptyset$, $i = 1$ ou 2 and with the original radiation condition Eq.(1.29). An overlap of $d > 0$ along x is actually considered to emphasize the differences between overlapping and non-overlapping methods. Hence

$$\Omega_1 =]-\infty, d[\times\mathbb{R} \quad \text{and} \quad \Omega_2 =]0, \infty[\times\mathbb{R} \tag{2.32}$$

and thus

$$\Gamma_{12} = \partial\Omega_1 \cap \overline{\Omega_2} = \{d\} \times \mathbb{R} \quad \text{and} \quad \Gamma_{21} = \partial\Omega_2 \cap \overline{\Omega_1} = \{0\} \times \mathbb{R} \tag{2.33}$$

are considered. From the intuitive interpretation of Jacobi iteration Eqs(2.27), the system to solve is

$$\begin{cases} \left(-\text{ div } \mathbf{grad} \ - k_0^2\right) u_1^{n+1} = f & \text{in } \Omega_1, \\ \lim\limits_{\rho\to\infty, x<0} \sqrt{\rho}\left(\dfrac{\partial}{\partial\rho} + ik_0\right) u_1^{n+1} = 0 & \\ \mathcal{J}_1 u_1^{n+1} = \mathcal{J}_1 u_2^n & \text{on } \Gamma_{12}, \end{cases} \tag{2.34}$$

$$\begin{cases} \left(-\text{ div } \mathbf{grad} \ - k_0^2\right) u_2^{n+1} = f & \text{in } \Omega_2, \\ \lim\limits_{\rho\to\infty, x>0} \sqrt{\rho}\left(\dfrac{\partial}{\partial\rho} + ik_0\right) u_2^{n+1} = 0 & \\ \mathcal{J}_2 u_2^{n+1} = \mathcal{J}_2 u_1^n & \text{on } \Gamma_{21}. \end{cases} \tag{2.35}$$

Subtracting the exact solution to the original (undecomposed) problem u^* to Eq.(2.34) and Eq.(2.35) yields[5], for $\{i, j\} = \{1, 2\}$ and $\{2, 1\}$,

$$\begin{cases} \left(-\text{ div } \mathbf{grad} \ - k_0^2\right) \epsilon_i^{n+1} = 0 & \text{in } \Omega_i, \\ \lim\limits_{\rho\to\infty} \sqrt{\rho}\left(\dfrac{\partial}{\partial\rho} + jk_0\right) \epsilon_i^{n+1} = 0, & \\ \mathcal{J}_i \epsilon_i^{n+1} = \mathcal{J}_i \epsilon_j^n & \text{on } \Gamma_{ij}, \end{cases} \tag{2.36}$$

where $\epsilon_i \triangleq u_i - u^*$ is the error on the solution given by the decomposition method. Eqs.(2.36) is therefore called the *error equation*. Applying the exact same treatment as in Appendix B.1, *i.e.* taking the Fourier transform w.r.t. y and solving the resulting ordinary differential equation, one obtains

$$\hat{\epsilon}_1^{n+1} = A_1^{n+1}(k_y) \exp \lambda x \tag{2.37}$$

$$\hat{\epsilon}_2^{n+1} = B_2^{n+1}(k_y) \exp -\lambda x \tag{2.38}$$

where A and B are constant w.r.t. x that can be determined using the transmission conditions.

Convergence rate A scheme is said to be convergent if the error decreases to zero, *i.e.* if $\forall x, k_y \in \mathbb{R}$,

$$\lim_{n'\to\infty} \left|\hat{\epsilon}_i^{n'}\right| = 0 \tag{2.39}$$

which is equivalent to

$$\lim_{n\to\infty} \left|\frac{\hat{\epsilon}_i^{2n}}{\hat{\epsilon}_i^{2(n-1)}}\right| \left|\frac{\hat{\epsilon}_i^{2(n-1)}}{\hat{\epsilon}_i^{2(n-2)}}\right| \cdots \left|\frac{\hat{\epsilon}_i^{2}}{\hat{\epsilon}_i^{0}}\right| \left|\hat{\epsilon}_i^{0}\right| = 0. \tag{2.40}$$

[5] As $\mathcal{J}_i$ is linear, $\mathcal{J}_i(u_i - u) = \mathcal{J}_i(\epsilon_i) = \mathcal{J}_i(u_i) - \mathcal{J}_i(u)$

Defining the convergence rate[6]

$$\rho(k_y) \triangleq \frac{\hat{\epsilon}_i^{n+1}}{\hat{\epsilon}_i^{n-1}} \left(= \frac{(A \text{ or } B)_i^{n+1}}{(A \text{ or } B)_i^{n-1}} \right), \tag{2.41}$$

the convergence is ensured for modes (*i.e.* values of k_y) such that

$$\lim_{n\to\infty} \left| \frac{\hat{\epsilon}_i^{2n}}{\hat{\epsilon}_i^{2(n-1)}} \right| \left| \frac{\hat{\epsilon}_i^{2(n-1)}}{\hat{\epsilon}_i^{2(n-2)}} \right| \dots \left| \frac{\hat{\epsilon}_i^{2}}{\hat{\epsilon}_i^{0}} \right| \left| \hat{\epsilon}_i^0 \right| = \lim_{n\to\infty} |\rho^n| \left| \hat{\epsilon}_i^0 \right| = 0, \tag{2.42}$$

i.e. such that

$$\boxed{|\rho(k_y)| < 1.} \tag{2.43}$$

Moreover, the smaller the convergence rate, the faster the convergence.

2.4.1 Classical Schwarz operator

The classical Schwarz method [7] suggests to use the simple boundary condition

$$\mathcal{J}_i = \mathcal{I}, \tag{2.44}$$

$\mathcal{I}$ being the identity operator. Using this operator in the boundary condition of Eqs(2.36) yields

$$A_1^{n+1} \exp \lambda d = B_2^n \exp -\lambda d \tag{2.45}$$
$$B_2^{n+1} = A_1^n \tag{2.46}$$

or

$$A_1^{n+1} = A_1^{n-1} \exp -2\lambda d \tag{2.47}$$
$$B_2^{n+1} = B_2^{n-1} \exp -2\lambda d \tag{2.48}$$

and the convergence rate is thus given for real wavenumber by

$$\rho(k_y) = \begin{cases} \exp\left(-2\sqrt{k_y^2 - k_0^2}d\right) & \text{for } |k_y| \geq |k_0| \\ \exp\left(-2j\sqrt{k_0^2 - k_y^2}d\right) & \text{for } |k_y| \leq |k_0|\,. \end{cases} \tag{2.49}$$

This convergence rate is represented in blue in Figure 2.1. Propagative (*i.e.* $|k_y| < |k_0|$) and resonant (*i.e.* $|k_y| = |k_0|$) modes never converge while evanescent modes (*i.e.* $k_y > k_0$) only converge if there is overlap (*i.e.* $d \neq 0$). This method is therefore not suited for non-overlapping domain decomposition. This transmission condition is referred to as CS.

[6]The convergence rates for the considered transmission boundary conditions, model problem and iterative scheme are actually independent of n and x.

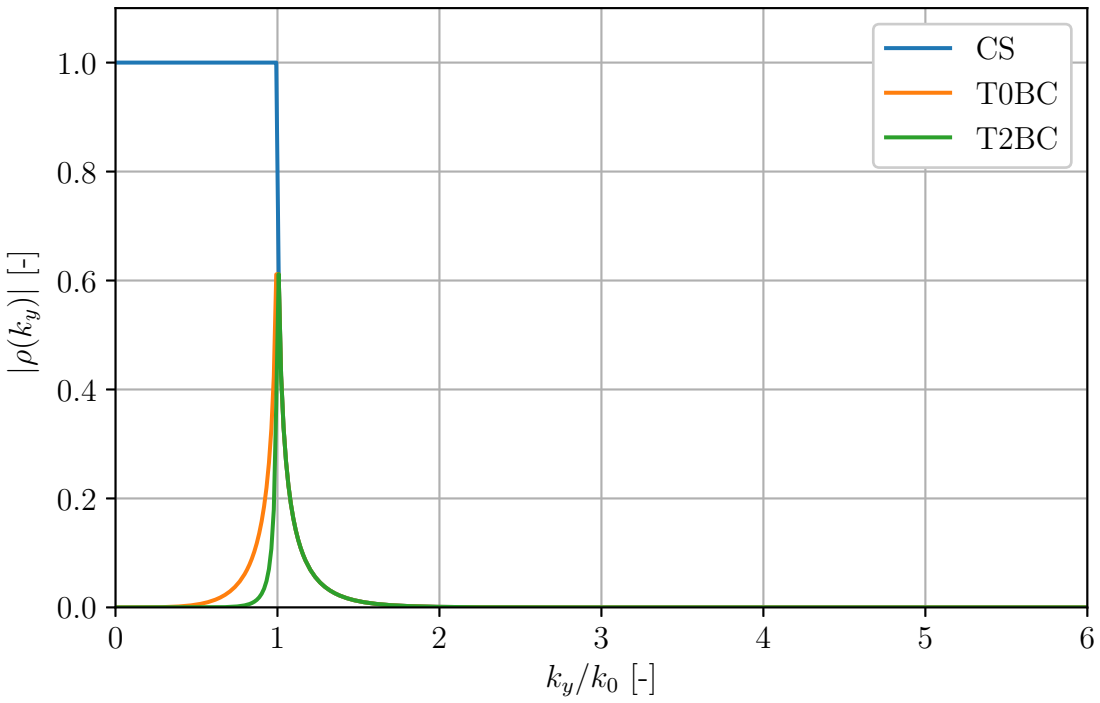


Figure 2.1: Convergence rates $|\rho(k_y)|$ for T0BC Eq.(2.59), T2BC Eq.(2.60) and CS Eq.(2.44) with overlap. $k_0 = 1$, $d = 2$.
This value of d typically corresponds to a case where the overlap is four finite elements. Indeed $k_0 = 1$ implies $\lambda = 2\pi$. As the characteristic mesh size h is usually taken as $h \approx \lambda/12$, $d = 4h \approx 2$.

2.4.2 Dirichlet-to-Neumann based operators

The family of transmission operators based on the DtN map considers operator $\mathcal{J}_i$ of the form

$$\mathcal{J}_i = \hat{\boldsymbol{n}}_i \cdot \mathbf{grad} \; + \mathcal{S}_i \tag{2.50}$$

where $\mathcal{S}_i$ is a linear operator acting only along the tangential direction of the interface and is an approximation of the exact DtN introduced in Appendix B.1. The Fourier transform w.r.t. y of the boundary condition of Eqs(2.36) then writes as

$$\left(\partial_x + \sigma_1\right) \hat{\epsilon}_1^{n+1} = \left(\partial_x + \sigma_1\right) \hat{\epsilon}_2^{n} \quad \text{for } x = d, k_y \in \mathbb{R}, \tag{2.51}$$
$$\left(\partial_x - \sigma_2\right) \hat{\epsilon}_2^{n+1} = \left(\partial_x - \sigma_2\right) \hat{\epsilon}_1^{n} \quad \text{for } x = 0, k_y \in \mathbb{R}, \tag{2.52}$$

where σ_i is the symbol of $\mathcal{S}_i$ (*i.e.* such that $\mathcal{F}\left(\mathcal{S}_i(\epsilon_i^{n+1})\right) = \sigma_1 \hat{\epsilon}_1^{n+1}$). These transmission boundary conditions then yield

$$\left(\lambda + \sigma_1\right) A^{n+1} \exp\left(\lambda d\right) = \left(-\lambda + \sigma_1\right) B^{n} \exp\left(-\lambda d\right), \tag{2.53}$$
$$\left(-\lambda - \sigma_2\right) B^{n+1} = \left(\lambda - \sigma_2\right) A^{n}, \tag{2.54}$$

which can be written as

$$A^{n+1} = \frac{-\lambda + \sigma_1}{\lambda + \sigma_1} \frac{\lambda - \sigma_2}{-\lambda - \sigma_2} \exp\left(-2\lambda d\right) A^{n-1} \tag{2.55}$$
$$B^{n+1} = \frac{-\lambda + \sigma_1}{\lambda + \sigma_1} \frac{\lambda - \sigma_2}{-\lambda - \sigma_2} \exp\left(-2\lambda d\right) B^{n-1} \tag{2.56}$$

and thus

$$\rho(k_y) = \frac{\sigma_1 - \lambda}{\sigma_1 + \lambda}\frac{\sigma_2 - \lambda}{\sigma_2 + \lambda}\exp\left(-2\lambda d\right). \tag{2.57}$$

Taking

$$\sigma_i = \lambda \tag{2.58}$$

which implies, by definition, that the transmission operator is the DtN map, *i.e.* $S_i = \Lambda$, cancels the convergence rate $\rho(k_y)\ \forall k_y$ and is thus the best choice one can find. Such a choice is however often avoided as the DtN operator is non local (*cfr* Section 1.6.1). As for absorbing boundary conditions, Taylor approximations of the DtN operator which are valid for waves of incidence close to the normal *i.e.* $k_y/k_0 \approx 0$, are considered. For the following derivations, the case were the wavenumber is real or complex have been separated. The most pathological case being the real one. Indeed when k has a complex part, dissipation is present and hence error also dissipates.

(a) Real wavenumber

The wavenumber can be considered real for so called bad conductors (*cfr* Section 1.1). For real wavenumbers, it is important to emphasize that the resonant mode, *i.e.* $k_y = k_0$, is never convergent for a transmission operator of the type Eq.(2.50) because $\rho(k_y = k_0) = 1$ (except if $\sigma_i = \lambda$).

Zeroth order Taylor approximation The zeroth order Taylor expansion of the DtN map is

$$\begin{aligned} & \sigma_i^{\text{T0BC}} = ik_0 \\ \Leftrightarrow \quad & \boxed{\mathcal{S}_i^{\text{T0BC}} = \text{Op}(\sigma_i^{\text{T0BC}}) = ik_0.} \end{aligned} \tag{2.59}$$

The related convergence rate is represented in Figure 2.1 with overlap and in Figure 2.2 without overlap. This transmission condition is referred to as T0BC.

Second order Taylor approximation The second order Taylor expansion of the DtN map is

$$\begin{aligned} & \sigma_i^{\text{T2BC}} = ik_0\left(1 - \frac{1}{2}\frac{k_y^2}{k_0^2}\right) \\ \Leftrightarrow \quad & \boxed{\mathcal{S}_i^{\text{T2BC}} = \text{Op}(\sigma_i^{\text{T2BC}}) = ik_0 - \frac{1}{2ik_0}\frac{\partial^2}{\partial^2 y}.} \end{aligned} \tag{2.60}$$

The related convergence rate is represented in Figure 2.1 with overlap and in Figure 2.2 without overlap. This transmission condition is referred to as T2BC.

As expected, without overlap, only modes with nearly normal incidence, *i.e.* $k_y/k_0 \approx 0$ converge, at the opposite of the classical Schwarz scheme. However, with overlap, Taylor approximations yield convergence for all propagative and evanescent modes. These results will not be discussed further as only non-overlapping domain decomposition methods are investigated.

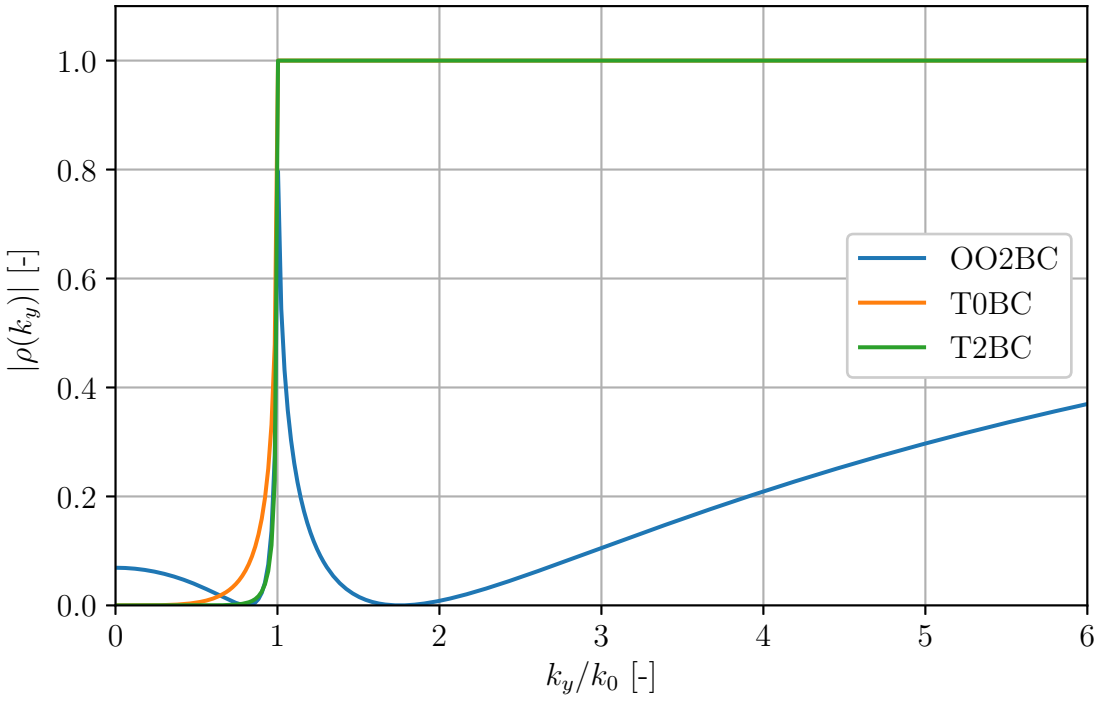


Figure 2.2: Convergence rates $|\rho(k_y)|$ for zeroth order Taylor approximation Eq.(2.59), second order Taylor approximation Eq.(2.60) and optimized second order Eq.(2.61) transmission conditions with overlap. $k_0 = 1$, $d = 0$, $k_{\text{max}} = 6$, $k_{\text{min}} = 6/100$, $k_- = k_0 - k_{\text{min}}$,$k_- = k_0 + k_{\text{min}}$. These values corresponds to a domain of 100 elements of characteristic size $h = \lambda/12$.

While Taylor approximations are attractive because they are easy to derive, they are however not well suited without overlap. Indeed, they cancel the convergence rate for low frequency modes ($k_y/k_0 \approx 0$) but not for all modes[7]. In the following paragraph, an optimal transmission condition is described.

Optimized second order approximation [11] A transmission condition performs well if the convergence rate is close to zero for all modes. [11] proposes to find the second order approximation that yields the lower convergence rate possible for the particular model problem developed in this section.
This *optimized second order approximation* transmission condition is given by

$$\mathcal{S}_i^{\text{002BC}} = \frac{\alpha^*\beta^* - k_0^2}{\alpha^* + \beta^*} - \frac{1}{\alpha^* + \beta^*}\frac{\partial^2}{\partial^2 y} \tag{2.61}$$

with

$$\alpha^* = i\left((k_0^2 - k_{\text{min}}^2)(k_0^2 - k_-^2)\right)^{1/4} \tag{2.62}$$

and

$$\beta^* = \left((k_{\text{max}}^2 - k_0^2)(k_+^2 - k_0^2)\right)^{1/4}. \tag{2.63}$$

[7] An intuitive interpretation is that they only transmit waves with nearly normal incidence because waves with non normal incidence ($k_y/k_0 >> 0$) are reflected, as explained in section 1.6 for absorbing boundary conditions. This reflection once again comes from the bad approximation of the DtN operator.

The parameters $k_{\min}$, $k_{\max}$, k_+ and k_- are problem dependant. $k_{\max}$ is the largest wave number representable on the numerical grid. If h is the characteristic mesh size, then a good choice is $k_{\max} = \pi/h$. Similarly for the minimum wave number $k_{\min}$, if the domain size is L, then a good choice is $k_{\min} = \pi/L$. The parameters k_- and k_+ are actually two values that must straddle the resonant frequency k_0[8]. As explained in [11], taking $k_- = k_0 - k_{\min}$ and $k_+ = k_0 + k_{\min}$ is a good choice because only the resonant frequency is rejected from the minimization of the convergence rate.
The convergence rates relative to this optimized transmission condition is represented in Figure 2.2. As expected the convergence is ensured for all modes except the resonant one. This transmission condition is referred to as OO2BC.

Other transmission conditions such as [4] can also be chosen.

(b) Complex wavenumber

When the wavenumber k is complex *i.e.* when the medium is either a good conductor or a dielectric, then dissipation is present (*cfr* Section 1.1). Results on the convergence rate are still valid but the approximations slightly differ. Indeed, in such a case, the principal symbol of the DtN becomes

$$\begin{aligned}\lambda &= \sqrt{k_y^2 - k^2} \\ &= \sqrt{k_y^2 - \omega^2\tilde{\epsilon}\mu_0} \\ &= \sqrt{k_y^2 - k_0^2 + i\omega\tilde{\sigma}\mu_0}.\end{aligned} \tag{2.64}$$

An important difference w.r.t. the real case is that the convergence rate Eq.(2.57) is not of modulus one for any transmission condition (*i.e.* for any value of σ_i) at the resonant mode $k_y = k_0$.

Good conductor Using the good conductor approximation, the principal symbol of the DtN then becomes

$$\lambda \approx \sqrt{k_y^2 + i\omega\tilde{\sigma}\mu_0} \tag{2.65}$$

such that the zeroth order Taylor approximation of this symbol gives

$$\sigma_i^{\text{T0GC}} \approx (1+i)\sqrt{\frac{\omega\tilde{\sigma}\mu_0}{2}} = \frac{(1+i)}{\delta} \tag{2.66}$$

and the second order Taylor approximation is

$$\sigma_i^{\text{T2GC}} \approx \frac{1+i}{\delta}\left(1 + \frac{k_y^2\delta^2}{4i}\right) \tag{2.67}$$

where $\delta = \sqrt{\frac{2}{\omega\tilde{\sigma}\mu_0}}$ [m] is the *skin depth*, a well-known parameter in electromagnetic theory. This parameter is the characteristic distance that wave can travel inside a good conductor before undergoing a strong attenuation (*i.e.* as $\exp\left(-\frac{x}{\delta}\right)$).
Both related convergence rates are represented in Figure 2.3 for different values of the ratio $\frac{\tilde{\sigma}}{\omega\text{Re}(\epsilon)}$. These transmission conditions are referred to as T0GC and T2GC.

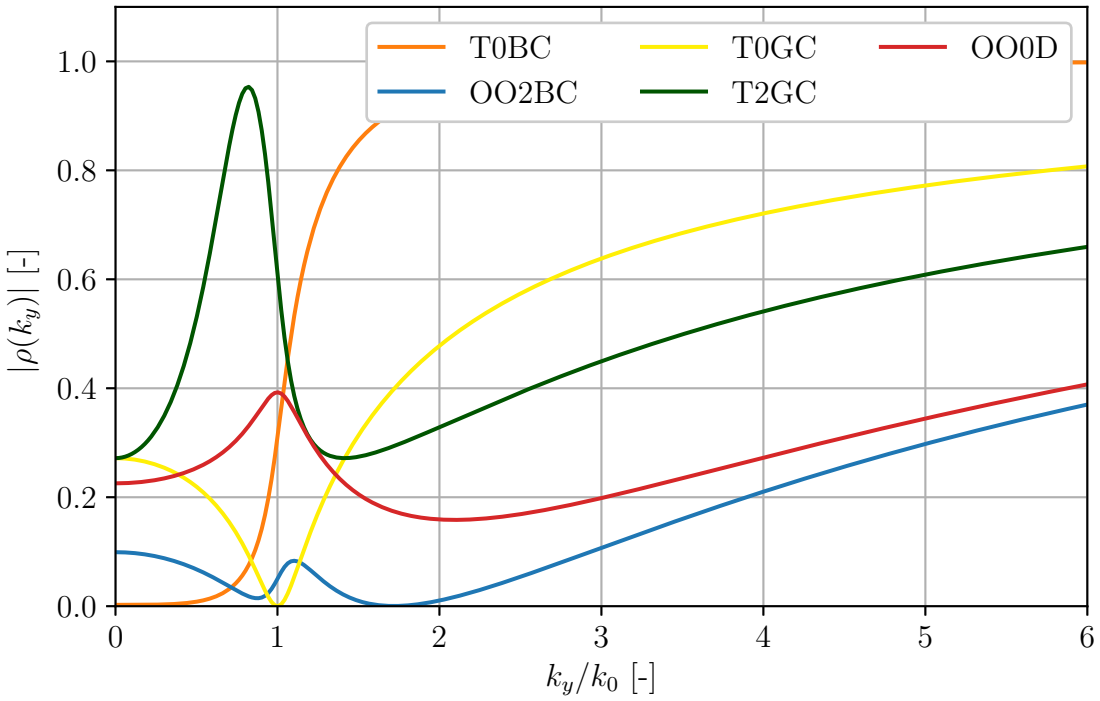
[8]This values are actually important in [11] because for a second order approximation, the resonant modes ($k_y = k_0$) is never convergent and this point must thus to be rejected from the minimization problem.

Optimized zeroth order transmission condition As previously, it is possible to find the best transmission conditions such that the convergence rate is minimal, for the range of consider wavenumbers $[k_{\text{min}}, k_{\text{max}}]$. In [6], it is proved that this optimal transmission condition is given by

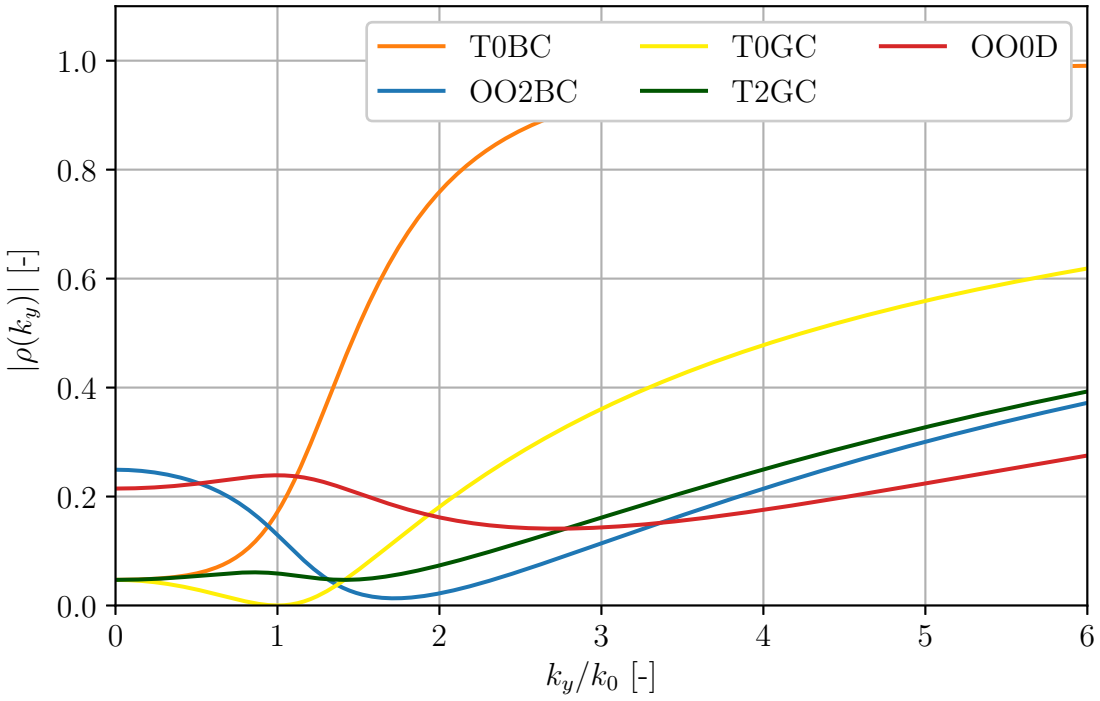
$$\boxed{\mathcal{S}_i^{\text{OO0D}} = (1+i)\sqrt{\frac{k_{\text{max}}}{\delta}}} \tag{2.68}$$

which is very similar to Eq.(2.66). The only difference is actually that Eq.(2.68) minimizes the convergence rate $\forall k_y \in [k_{\text{min}}, k_{\text{max}}]$ while Eq.(2.66) minimizes the convergence rate only when $k_y \ll k_0$. The related convergence rate is represented in Figure 2.3 for different value of the ratio $\frac{\tilde{\sigma}}{\omega \text{Re}(\epsilon)}$.
This transmission condition is referred to as TOO0D.

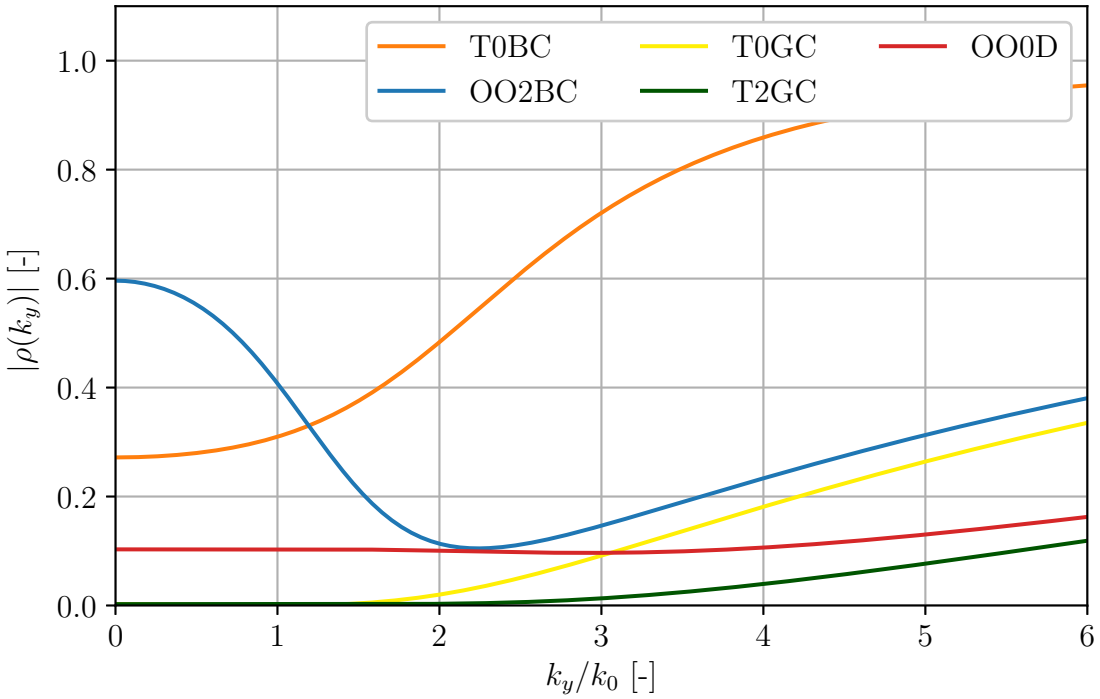
From Figure 2.3, the smoothing of the resonant mode (*i.e.*$k_y = k_0$) provided by dissipation can be observed, in particular for T0BC and OO2BC where it strongly differs from the case real case.
As expected, T0GC and T2GC only perform well when the hypotheses used to derive the transmission conditions are satisfied, *i.e.* when $k_y \ll k_0$ and $\frac{\tilde{\sigma}}{\omega \text{Re}(\epsilon)} \gg 1$.
OO0D is the best transmission condition when looking at all the possible media. However for dielectric (*i.e.* $\frac{\tilde{\sigma}}{\omega \text{Re}(\epsilon)} \approx 1$) and bad conductors (*i.e.* $\frac{\tilde{\sigma}}{\omega \text{Re}(\epsilon)} \ll 1$), OO2BC has also a small convergence rate. The choice between OO2D and OO2BC will therefore depend on the media involved.



(a) $\frac{\tilde{\sigma}}{\omega \mathrm{Re}(\epsilon)} = 0.2$



(b) $\frac{\tilde{\sigma}}{\omega \mathrm{Re}(\epsilon)} = 1$



(c) $\dfrac{\tilde{\sigma}}{\omega \mathrm{Re}(\epsilon)} = 5$

Figure 2.3: Convergence rates $|\rho(k_y)|$ for T0BC Eq.(2.59), OO2BC Eq.(2.61), T0GC Eq.(2.66) and OO0D Eq.(2.68). No overlap is considered. $k_0 = 1$, $d = 0$, $k_{\text{max}} = 6$, $k_{\text{min}} = 6/100$, $k_- = k_0 - k_{\text{min}}$,$k_- = k_0 + k_{\text{min}}$. These values correspond to a domain of 100 elements of characteristic size $h = \lambda/12$.

Chapter 3

Nodal finite element discretization

The two previous chapters introduced some non-overlapping Schwarz algorithms considering continuous unknowns.
To implement these algorithms, one should be able to compute the action of $\boldsymbol{\mathcal{A}}$ on any input artificial sources/surface unknown and to compute $\boldsymbol{b}$. By definition this is equivalent to solve the volume problem Eqs(2.7) and Eqs(2.9) and to apply the surface equations Eq.(2.6) and Eq.(2.8).
In this chapter only the system for $\boldsymbol{\mathcal{A}}$ will be considered as both systems for $\boldsymbol{\mathcal{A}'}$ and $\boldsymbol{b}$ are particular case of it (as emphasized by Eq.(2.10) and Eq.(2.11)). The solution to this system is obtain using the finite element method. Other numerical methods such as finite differences could be chosen as well.

3.1 Weak forms

The finite element method is not based on the *strong form* of the system that was used up to here but rather on a so called *weak form*. This weak form is obtained by multiplying the equations to verify by a test function $\tilde{w}$ or $\tilde{\mathcal{A}}_{ji}$ then by integrating over the space where the equations is verified. It can be shown that the weak and the strong form are equivalent [1]. The weak form relative to the volume problem is

$$\begin{aligned}\int_{\Omega_i} \mathbf{grad}\ \tilde{w} \cdot \mathbf{grad}\ w \, d\Omega_i - \omega^2 \int_{\Omega_i} \left(\frac{k}{\omega}\right)^2 \tilde{w} w \, d\Omega_i \\ + \int_{\partial\Omega_i^\infty} \tilde{w} \mathcal{B} w \, d\partial\Omega_i^\infty + \sum_{m \in D_i} \int_{\Gamma_{im}} \tilde{w} \mathcal{S} w \, d\Gamma_{im} \\ = \int_{\Omega_i} \tilde{w} f \, d\Omega_i + \int_{\partial\Omega_i^\infty} \tilde{w} u_\infty \, d\partial\Omega_i^\infty + \int_{\partial\Omega_i^n} \tilde{w} u_n \, d\partial\Omega_i^\infty + \sum_{m \in D_i} \int_{\Gamma_{im}} \tilde{w} s_{im} \, d\Gamma_{im}.\end{aligned} \tag{3.1}$$

The organization of Eq.(3.1) is not random: sources terms are on the right hand side while unknowns are on the left hand side. The weak form relative to the surface equation is

$$\int_{\Gamma_{ij}} \tilde{\mathcal{A}}_{ji} \mathcal{A}_{ji} \, d\Gamma_{ij} = -\int_{\Gamma_{ij}} \tilde{\mathcal{A}}_{ji} s_{ij} \, d\Gamma_{ij} + \int_{\Gamma_{ij}} \tilde{\mathcal{A}}_{ji} \left(\mathcal{S}_i + \mathcal{S}_j\right) w \, d\Gamma_{ij}. \tag{3.2}$$

Once again, the unknown functions are on the left hand side and the known functions are on the right hand side. Indeed this equation is only used when w is known. Derivations details for these two equations are given in Appendix B.4.

In the following sections, the boundary and transmission condition operators are chosen to be of the form

$$\mathcal{B} = a_\infty - b_\infty \mathrm{div}_{\partial\Omega^\infty}\, \mathbf{grad}_{\partial\Omega^\infty} \quad \text{and} \quad \mathcal{S} = a_\Gamma - b_\Gamma \mathrm{div}_\Gamma\, \mathbf{grad}_\Gamma \tag{3.3}$$

as derived in Section 1.6 and 2.4.

3.2 Discretization

To be equivalent to the strong forms, Eq.(3.1) (resp. Eq.(3.2)) must be verified for any test function $\tilde{w}$ (resp. $\tilde{\mathcal{A}}_{ji}$). The number of possibilities is infinite and thus impossible to verify in practice. The idea of finite elements is then to restrict the space to which w and $\tilde{w}$ (resp. $\tilde{\mathcal{A}}_{ji}$ and $\mathcal{A}_{ji}$) belong, to a finite dimension space such that Eq.(3.1) (resp. Eq.(3.2)) becomes discrete and can then be solved numerically.
A brief description of the steps required to obtain these discrete equations is given in the following paragraphs. More complete treatment can be found in litterature, *e.g.* [1].

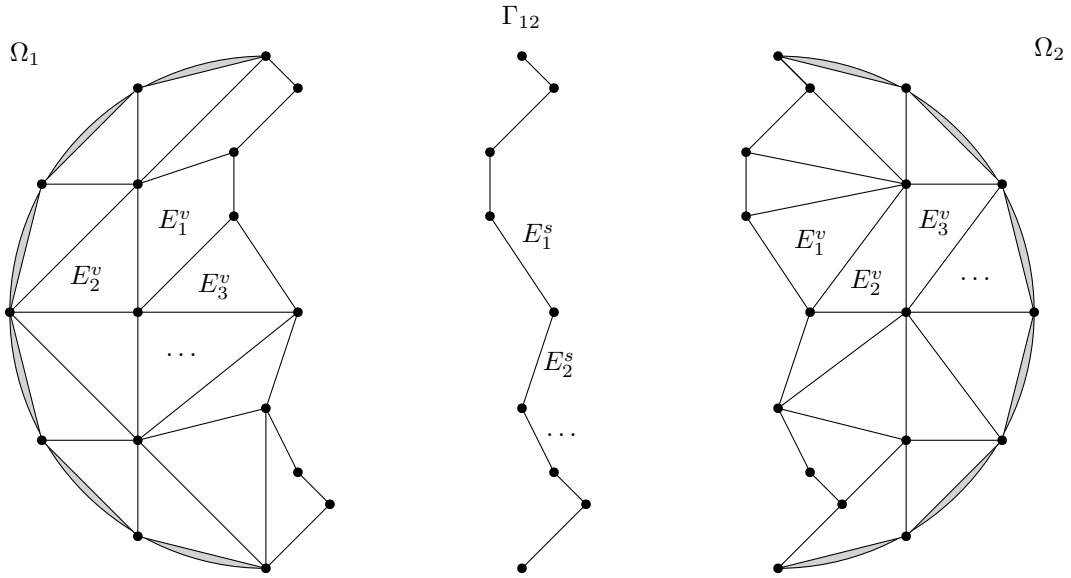


Figure 3.1: Finite element mesh example in the case of a disc divided in two subdomains and their common boundary.

Mesh generation The first step of a finite element algorithm is to decompose the resolution domain Ω_i (resp. Γ_{ij}) into smaller simple volumes, *e.g.* triangles or quadrangles (resp. line segments). These smaller volumes, denoted by E_ν^v (resp. E_ν^s) and called *elements*, should approximate the true resolution domain as well as possible, that is

$$\bigcup_\nu E_\nu \approx \Omega_i \quad \left(\text{resp. } \bigcup_\nu E_\nu \approx \Gamma_{ij} \right). \tag{3.4}$$

In the scope of domain decomposition, the mesh must also be splitted. An example of a splitted mesh is given in Figure 3.1 for a disc.

Node projection The mesh is the basis for the projection of w and $\tilde{w}$ (resp. $\tilde{\mathcal{A}}_{ji}$ and $\mathcal{A}_{ji}$) into a finite dimension space. Indeed, in a finite element approximation, these fields are expressed as a sum of elementary functions related to elementary entities of the elements. Depending on the problem, these elementary entities can be nodes, edges or surfaces of the elements.
For the scalar two dimensional case that is considered in this work, the fields are expressed as a sum of *hat node functions* ψ_k. These functions are denoted by ψ_k^v for fields that live on Ω_i (*i.e.* w and $\tilde{w}$) and by ψ_k^s for fields that live on $\bigcup_{j\in D_i}\Gamma_{ij}$ (*i.e.* $\tilde{\mathcal{A}}_{ji}$ and $\mathcal{A}_{ji}$). These hat functions are linear per element and such that

$$\psi_k^v(\boldsymbol{x}_i) = \begin{cases} 1 & i = k, \\ 0 & i \neq k \quad \forall k, i \in N_{\Omega_i} \end{cases} \tag{3.5}$$

or

$$\psi_k^s(\boldsymbol{x}_i) = \begin{cases} 1 & i = k, \\ 0 & i \neq k \quad \forall k, i \in N_{\Gamma_{ij}} \end{cases} \tag{3.6}$$

where N_{Ω_i} and $N_{\Gamma_{ij}}$ are respectively the set of nodes belonging to Ω_i and $\bigcup_{j\in D_i}\Gamma_{ij}$, and $\boldsymbol{x}_i$ is the position of nodes i.
Consequently, each function ψ_k has a compact support defined by the elements that contained node k. Other choices of node functions can be made but the compactness of their support over a few elements around node k must still be ensured. A schematic representation of a node functions is given in Figure 3.2.

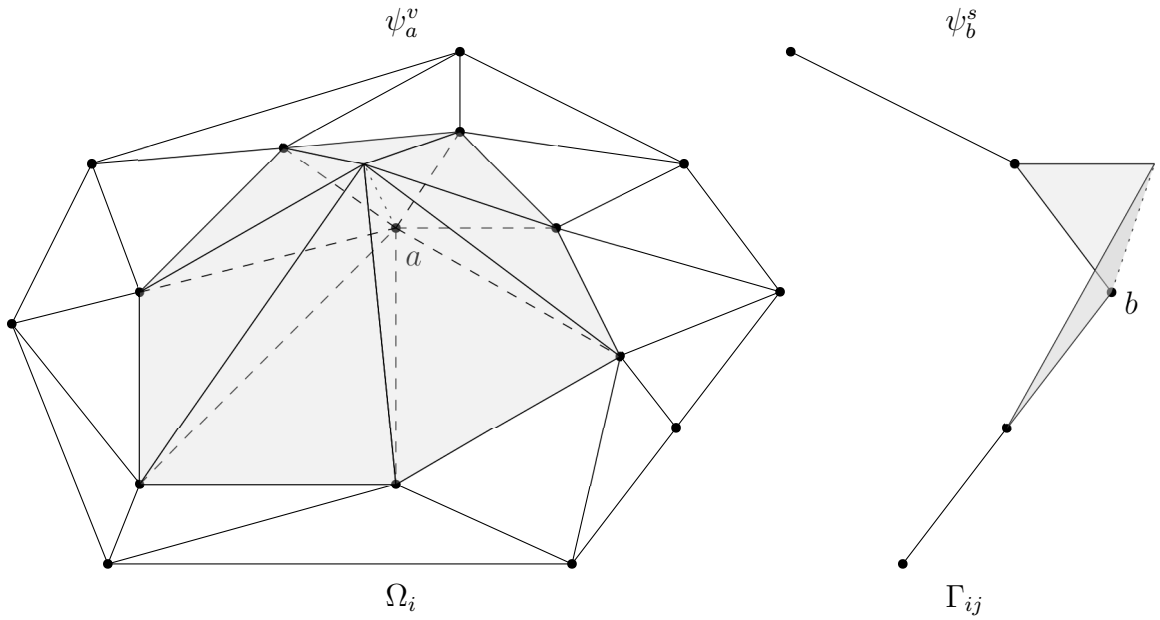


Figure 3.2: Representation of the hat node functions ψ_a^v associated to node a and ψ_b^s associated to node b.

Using this approximation, fields of the weak forms Eq.(3.1) and Eq.(3.2) then become

$$w \approx \sum_{k\in N_{\Omega_i}} w_k \psi_k^v, \quad \tilde{w} \approx \sum_{k\in N_{\Omega_i}} \tilde{w}_k \psi_k^v, \tag{3.7}$$

and

$$\mathcal{A}_{ji} \approx \sum_{k \in N_{\Gamma_{ij}}} \mathcal{A}_{ji,k} \psi_k^s, \quad \tilde{\mathcal{A}}_{ji} \approx \sum_{k \in N_{\Gamma_{ij}}} \tilde{\mathcal{A}}_{ji,k} \psi_k^s. \tag{3.8}$$

For the sake of simplicity, in the following paragraphs, the nodal values w_k, $\tilde{w}_k$, $\mathcal{A}_{ji,k}$, $\tilde{\mathcal{A}}_{ji,k}$ are gathered in vectors denoted by

$$\underline{w}, \quad \underline{\tilde{w}}, \quad \underline{\mathcal{A}}_{ij}, \quad \text{and} \quad \underline{\tilde{\mathcal{A}}}_{ij} \tag{3.9}$$

Discrete equations Inserting the approximations Eqs(3.7) and Eqs(3.8) into the weak forms Eq.(3.1) and Eq.(3.2) yields two discrete systems. The discrete system relative to the volume problem is

$$\begin{aligned} \underline{\tilde{w}}^T \underline{\underline{K}}_{\Omega_i}^{vv} \underline{w} - \omega^2 \underline{\tilde{w}}^T \underline{\underline{M}}_{\Omega_i}^{vv} \underline{w} + \left[a_\infty \underline{\tilde{w}}^T \underline{\underline{M}}_{\partial\Omega_i^\infty}^{vv} \underline{w} + b_\infty \underline{\tilde{w}}^T \underline{\underline{K}}_{\partial\Omega_i^\infty}^{vv} \underline{w} \right] + \sum_{m \in D_i} \left[a_\Gamma \underline{\tilde{w}}^T \underline{\underline{M}}_{\Gamma_{im}}^{vv} \underline{w} + b_\Gamma \underline{\tilde{w}}^T \underline{\underline{K}}_{\Gamma_{im}}^{vv} \underline{w} \right] \\ = \underline{\tilde{w}}^T \underline{f}_{\Omega_i} + \underline{\tilde{w}}^T \underline{u}_{\infty,\partial\Omega_i^\infty} + \underline{\tilde{w}}^T \underline{u}_{n,\partial\Omega_i^n} + \sum_{m \in D_i} \underline{\tilde{w}}^T \underline{\underline{M}}_{\Gamma_{im}}^{vs} \underline{s}_{im} \end{aligned} \tag{3.10}$$

while the discrete system relative to the surface equation is

$$\underline{\tilde{\mathcal{A}}}_{ji}^T \underline{\underline{M}}_{\Gamma_{ij}}^{ss} \underline{\mathcal{A}}_{ji} = -\underline{\tilde{\mathcal{A}}}_{ji}^T \underline{\underline{M}}_{\Gamma_{ij}}^{ss} \underline{s_{ij}} + 2 \underline{\tilde{\mathcal{A}}}_{ji}^T \left(a_\Gamma \underline{\underline{M}}_{\Gamma_{ij}}^{sv} + b_\Gamma \underline{\underline{K}}_{\Gamma_{ij}}^{sv} \right) \underline{w}. \tag{3.11}$$

$\underline{\underline{K}}$ and $\underline{\underline{M}}$ stand respectively for the *stiffness* and *mass* matrices. The mass matrix typically involves the nodes functions while stiffness matrices involve the gradient of these node functions. Explicit definitions of these matrices are given in Appendix B.5. As these equations must be verified for all $\underline{\tilde{w}}$ and $\underline{\tilde{\mathcal{A}}}_{ji}$, the systems to solve are then

$$\underline{\underline{K}}_i^t \underline{w} = \underline{f}_i^t + \sum_{m \in D_i} \underline{\underline{M}}_{\Gamma_{im}}^{vs} \underline{s}_{im} \tag{3.12}$$

and

$$\underline{\underline{M}}_{\Gamma_{ij}}^{ss} \underline{\mathcal{A}_{ji}} = -\underline{\underline{M}}_{\Gamma_{ij}}^{ss} \underline{s_{ij}} + 2 \left(a_\Gamma \underline{\underline{M}}_{\Gamma_{ij}}^{sv} + b_\Gamma \underline{\underline{K}}_{\Gamma_{ij}}^{sv} \right) \underline{w}, \tag{3.13}$$

where the *total stiffness matrix* is defined by

$$\underline{\underline{K}}_i^t \triangleq \underline{\underline{K}}_{\Omega_i}^{vv} - \omega^2 \underline{\underline{M}}_{\Omega_i}^{vv} + \left[a_\infty \underline{\underline{M}}_{\partial\Omega_i^\infty}^{vv} + b_\infty \underline{\underline{K}}_{\partial\Omega_i^\infty}^{vv} \right] + \sum_{m \in D_i} \left[a_\Gamma \underline{\underline{M}}_{\Gamma_{im}}^{vv} + b_\Gamma \underline{\underline{K}}_{\Gamma_{im}}^{vv} \right] \tag{3.14}$$

and the *total physical source* term is defined by

$$\underline{f}_i^t \triangleq \underline{f}_{\Omega_i} + \underline{u}_{\infty,\partial\Omega_i^\infty} + \underline{u}_{n,\partial\Omega_i^n}. \tag{3.15}$$

Combining Eq.(3.12) and Eq.(3.13) finally yields the finite element discretisation of the operator $\boldsymbol{\mathcal{A}}$

$$\boxed{\begin{aligned} \underline{\mathcal{A}}_{ji} = -\underline{s}_{ij} + \sum_{m \in D_i} 2 \underline{\underline{M}}_{\Gamma_{ij}}^{ss\ -1} \left(a_\Gamma \underline{\underline{M}}_{\Gamma_{ij}}^{sv} + b_\Gamma \underline{\underline{K}}_{\Gamma_{ij}}^{sv} \right) \underline{\underline{K}}_i^{t\,-1} \underline{\underline{M}}_{\Gamma_{im}}^{vs} \underline{s}_{im} \\ + 2 \underline{\underline{M}}_{\Gamma_{ij}}^{ss\ -1} \left(a_\Gamma \underline{\underline{M}}_{\Gamma_{ij}}^{sv} + b_\Gamma \underline{\underline{K}}_{\Gamma_{ij}}^{sv} \right) \underline{\underline{K}}_i^{t\,-1} \underline{f}_i^t. \end{aligned}} \tag{3.16}$$

The solution $\underline{w}$ to the volume problem is computed as the exact, direct solution of discrete system Eq.(3.12) while the solution to the surface problem Eq.(2.14) is computed iteratively using the discretization Eq.(3.13) of the $\boldsymbol{\mathcal{A}}$ operator. For this reason, the decomposition algorithm presented in

this work is qualified as an *hybrid solver*.
Each time the action of $\boldsymbol{\mathcal{A}'}$ on an input $\boldsymbol{s}$ of artificial sources/surface unknowns is computed during Jacobi or Krylov iterations, first Eq.(3.12) is solved for each subdomain, then Eq.(3.13) is computed for each subdomain too. One advantage of this formulation is that both Eq.(3.12) and Eq.(3.13) are independent for any subdomain provided $\boldsymbol{s}$ is known. The overall solution can therefore be computed in parallel very easily. Between two iterations, only the artificial sources $\boldsymbol{s}$, to be used for the next iterations, must be shared. It is important to point out that the discretization of $\boldsymbol{\mathcal{A}'}$ Eq.(3.16) is actually not computed explicitly as it requires the knowledge of $\underline{\underline{K}}_i^{t-1}$ which is never computed explicitly, only the solution of Eq.(3.12) is actually computed. More details about the discrete finite elements equations of Schwarz algorithm can be found in [11].

Matrix assembly Several mass and stiffness matrices have been introduced in the previous paragraph and their definition is given in Appendix B.5. These matrices are however not computed directly from their definition. Actually, because of the compact support of the hat node function, these matrices have only very few non zero entry. Consider for example

$$\left[\underline{\underline{M}}^{vv}_{\Omega_i}\right]_{pq} \triangleq \int_{\Omega_i} \psi_p^v \psi_q^v \, d\Omega_i, \quad \forall p \in N_{\Omega_i}, q \in N_{\Omega_i}. \tag{3.17}$$

The first step to compute this matrix is to use the geometrical finite elements approximation Eq.(3.4), *i.e.*

$$\left[\underline{\underline{M}}^{vv}_{\Omega_i}\right]_{pq} \approx \sum_{\nu \in E_{\Omega_i}} \int_{E_\nu} \psi_p^v \psi_q^v \, dE_\nu, \quad \forall p \in N_{\Omega_i}, q \in N_{\Omega_i}, \tag{3.18}$$

where E_{Ω_i} is the set of all elements of Ω_i.
Then because ψ_p^v (resp. ψ_q^v) vanishes everywhere outside elements that contain node p (resp. node q), the term

$$\int_{E_\nu} \psi_r \psi_s \, dE_\nu \tag{3.19}$$

is non vanishing only if $p \in N_{E_\nu}$ and $q \in N_{E_\nu}$. Thus the matrix is computed as

$$\left[\underline{\underline{M}}^{vv}_{\Omega_i}\right]_{pq} \approx \sum_{\nu \in E_{\Omega_i}} \int_{E_\nu} \psi_p^v \psi_q^v \, dE_\nu, \quad \forall p \in N_{E_\nu}, q \in N_{E_\nu}. \tag{3.20}$$

The matrices involved in the preceding paragraphs are sparse because the hat node function have a very compact support, increasing the support of these functions would decrease the sparsity of the matrix. Matrix assembly can actually be further simplified by introducing reference elements [19]. All the stiffness and mass matrices actually depend on the geometry, the mesh and its decomposition and the node functions but not on the artificial sources $\boldsymbol{s}$ nor on the physical sources f. Therefore these matrices are computed only once and used for any artificial or physical source.

Part II

Phased array optimization in heterogeneous media

Chapter 4

Phased arrays

In Part I, a solution technique for Maxwell's equations using domain decomposition for transverse electric or magnetic wave propagation problems is derived. In this chapter, this solution technique is applied to the simulation of phased arrays in heterogeneous media. Phased arrays are commonly used in antenna design [21] as they allow to achieve high and tunable directivity. This chapter first introduced the most general case of phased arrays then only a simpler case is investigated further. In particular, asymptotic results in homogeneous medium are given to gain intuition on how phased arrays work.

4.1 Definition

A phased array is, by definition, a particular distribution of sources that can be written

$$\boldsymbol{j}_a = \sum_{p=1}^{N_p} j_p \boldsymbol{w}_p \tag{4.1}$$

where $\boldsymbol{w}_p$ [-] are the *elementary source functions*, $j_p \in \mathbb{C}$ [A/m^2] are their weights and N_p is the number of elementary sources of the array. In practice, the elementary source functions are often taken to be the same function $\boldsymbol{w}$ shifted by $\boldsymbol{x}_p$, *i.e.*

$$\boldsymbol{w}_p = \boldsymbol{w}\left(\boldsymbol{x} - \boldsymbol{x}_p\right). \tag{4.2}$$

For simplicity, only transverse electric (TE) fields are considered[1]. Using the notations of Chapter 1 (in particular Eq.(1.22)), the source term f is given by

$$f = -i\omega\mu_0 \sum_{p=1}^{N_p} j_p w_p. \tag{4.3}$$

The Helmholtz equation and the Sommerfeld radiation condition Eqs(1.31) being linear, the solution for an array of sources can be written by superposition as

$$u = \sum_{p=1}^{N_p} j_p u_p \tag{4.4}$$

[1]For transverse magnetic (TM) fields, **curl** $\boldsymbol{j}_a = f(x,y)\hat{\boldsymbol{z}}$ (*cfr* Section 1.1) such that $\boldsymbol{j}_a \cdot \hat{\boldsymbol{z}} = 0$ and thus $\boldsymbol{w} \cdot \hat{\boldsymbol{z}} = 0$.

where u_p is the p^{th} *elementary solution*, *i.e.* it verifies[2]

$$\boxed{\begin{cases} \left(-\text{div } \mathbf{grad } - k^2\right) u_p = -i\omega\mu_0 j_0 w_p \quad \forall \boldsymbol{x} \in \mathbb{R}^2, \\ \lim\limits_{\rho\to\infty} \sqrt{\rho}\left(\dfrac{\partial u_p}{\partial \rho} + ik_0 u\right) = 0. \end{cases}} \tag{4.5}$$

4.2 Point source arrays

In this section, the elementary sources are taken to be *Dirac distributions*, *i.e.*

$$w_p(\boldsymbol{x}) = A_\delta \delta\left(\boldsymbol{x} - \boldsymbol{x}_p\right) \text{ with } A_\delta = 1 \text{ [m}^2\text{]}. \tag{4.6}$$

This kind of sources is chosen to be able to conduct simple analytic studies but also because it has applications[3]outside the scope of antennas and because several real elementary sources are equivalent to point source under some assumptions (*cfr* Section 4.3). The physical interpretation of this source is a wire carrying a constant current and considered infinitely long along the $\hat{\boldsymbol{z}}$-axis (such that the problem remains two dimensional) and infinitely thin in the $\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}$ plane. Using only the q^{th} source, the q^{th} elementary solution is then called a *Green function*[3] g and is defined as

$$\begin{aligned} g\left(\boldsymbol{x}; \boldsymbol{x}_q\right) \triangleq u_q\left(\boldsymbol{x}\right) \text{ [-]} \quad & \text{for } w_q(\boldsymbol{x}) = A_\delta \delta\left(\boldsymbol{x} - \boldsymbol{x}_q\right) \\ & \text{and } -i\omega\mu_0 j_p A_\delta \to \delta_{pq} \text{ [-]}, \forall p, q \in [1, N_p]. \end{aligned} \tag{4.7}$$

Free space Green function In the particular case where the medium is homogeneous, *i.e.* when $k(\boldsymbol{x}) \to k_0$, the *free space Green function* $g_0(\boldsymbol{x}, \boldsymbol{x}_0)$ is given by[4][5]

$$g_0(\boldsymbol{x}, \boldsymbol{x}_0) = -\frac{i}{4} H_0^{(2)}\left(k_0 \left\|\boldsymbol{x} - \boldsymbol{x}_0\right\|\right) \tag{4.9}$$

where $H_\alpha^{(2)}$ is the Hankel function of the second kind of order α, defined as $H_\alpha^{(2)} \triangleq J_\alpha - iY_\alpha$, J_α and Y_α being Bessel functions of the first and second kind of order α. A common approximation of the Hankel function of the second kind is [12]

$$H_0^{(2)}(z) \approx \sqrt{\frac{2}{\pi z}} \exp\left(-i\left(z - \pi/4\right)\right). \tag{4.10}$$

[2] $j_0 = 1$ [A/m^2].

[3] The complete knowledge of the Green function (*i.e.* for any position $\boldsymbol{y}$ of the Dirac distribution) is actually very useful as it can be shown under some assumptions that the general solution for any source distribution f is then

$$u(\boldsymbol{x}) = \int g(\boldsymbol{x}, \boldsymbol{y}) f(\boldsymbol{y}) d\boldsymbol{y}. \tag{4.8}$$

This formula is a generalization of the superposition principle used to write Eq.(4.4).

[4] It is interesting to point out that the Hankel function of the first kind $H_0^{(1)} \triangleq J_0 + iY_0$ is also a solution to the Helmholtz equation alone but this solution does however not satisfy the Sommerfeld radiation condition and must thus be discarded. Intuitively, the situation is the same than in Section 1.5, the radiation condition prescribes solution that yields energy coming from infinity towards the source.

This approximation is especially valid when the *dimensionless distance* $z \triangleq k_0 \|\boldsymbol{x} - \boldsymbol{x}_0\|$ is large. Physically, it means that $x - x_0$ is much greater than the wavelength $\lambda \triangleq 2\pi/k_0$ *i.e. electrically far* from the source. Hence, Eq.(4.9) with approximation Eq.(4.10) is called the *far field* generated by the Dirac source and is denoted by g_0^{FF}. The Green function Eq.(4.9) and its far field approximation g_0^{FF} are represented in Figure 4.1.

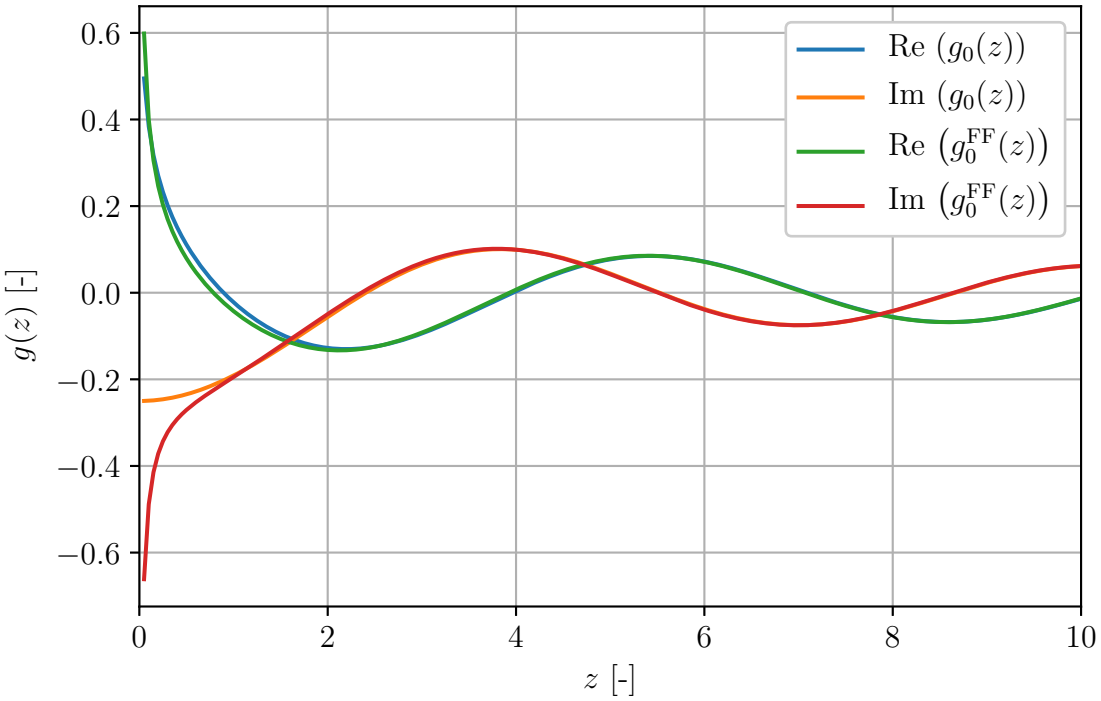


Figure 4.1: Exact and far field free space Green function g_0 and g_0^{FF} associated to the two dimensional scalar Helmholtz equation with Sommerfeld radiation condition. $z \triangleq k_0 \|\boldsymbol{x} - \boldsymbol{x}_0\|$.

4.2.1 Far field in free space

Consider a N_p-point source array, *i.e.*

$$f = -i\omega\mu_0 \sum_{p=1}^{N_p} j_p A_\delta \delta(\boldsymbol{x} - \boldsymbol{x}_p). \tag{4.11}$$

Using the superposition principle Eq.(4.4), the solution in free space then writes

$$u = -i\omega\mu_0 \sum_{p=1}^{N_p} j_p A_\delta g_0(\boldsymbol{x}, \boldsymbol{x}_p). \tag{4.12}$$

For the sake of simplicity, the weights $j_p \in \mathbb{C}$ are substituted by an amplitude $w_p \in \mathbb{R}^+$ and a phase $\alpha_p \in]-\pi, \pi]$ such that $-i\omega\mu_0 j_p A_\delta = w_p \exp(-i\alpha_p)$ [V/m] and the dimensionless distances $z_p \triangleq k_0 \|\boldsymbol{x} - \boldsymbol{x}_p\|$ are used.

More intuitive interpretation of this solution can be obtained by inserting the far field approximation Eq.(4.10) into Eq.(4.12), *i.e.*

$$\begin{aligned} u^{\text{FF}} &= \sum_{p=1}^{N_p} w_p \exp\left(-i\alpha_p\right) \frac{-i}{4} \sqrt{\frac{2}{\pi z_p}} \exp\left(i\pi/4\right) \exp\left(-iz_p\right) \\ &= \sum_{p=1}^{N_p} \frac{-i}{4} \sqrt{\frac{2}{\pi z_p}} \exp\left(i\pi/4\right) w_p \exp\left(-i(z_p+\alpha_p)\right). \end{aligned} \tag{4.13}$$

When several point sources are present, the far field approximations are

$$z_p \gg 1 \quad \forall p \in [1; N_p]. \tag{4.14}$$

If all the point sources are concentrated in a region of radius $x_{\max}$ $(= \mathcal{O}(\lambda))$ around the origin then the approximations reduces to

$$\|\boldsymbol{x}\| \gg x_{\max}. \tag{4.15}$$

Using a Taylor expansion on the definition of z_p then yields

$$\begin{aligned} z_p &\triangleq k_0 \sqrt{\|\boldsymbol{x}\|^2 + \|\boldsymbol{x}_p\|^2 - 2\boldsymbol{x}\cdot\boldsymbol{x}_p} \\ &= k_0 \|\boldsymbol{x}\| \sqrt{1 + \left(\frac{\|\boldsymbol{x}_p\|}{\|\boldsymbol{x}\|}\right)^2 - 2\frac{\boldsymbol{x}\cdot\boldsymbol{x}_p}{\|\boldsymbol{x}\|^2}} \\ &\approx k_0 \|\boldsymbol{x}\| \sqrt{1 - 2\frac{\boldsymbol{x}_p\cdot\boldsymbol{x}}{\|\boldsymbol{x}\|^2}} \\ &\approx k_0 \|\boldsymbol{x}\| \left(1 - \frac{\boldsymbol{x}\cdot\boldsymbol{x}_p}{\|\boldsymbol{x}\|^2}\right) \\ &= z\left(1 - \frac{\boldsymbol{x}\cdot\boldsymbol{x}_p}{\|\boldsymbol{x}\|^2}\right) \end{aligned} \tag{4.16}$$

Physically, this approximation consists in considering that $\boldsymbol{x} - \boldsymbol{x}_p$ and $\boldsymbol{x}$ are parallel, as represented schematically in Figure 4.2 for two point sources. In electromagnetic theory, this approximation is called the *Fraunhofer approximation* [21].

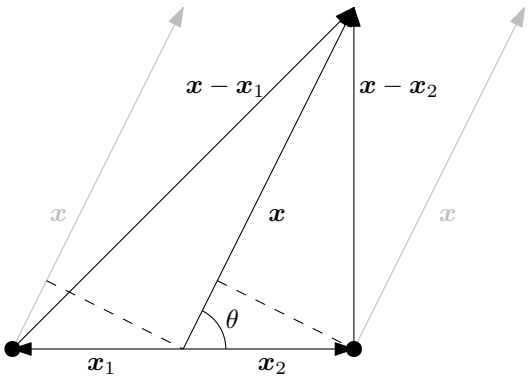


Figure 4.2: Fraunhofer approximation for two point sources.

Using the Fraunhofer approximation Eq.(4.16), the far field solution Eq.(4.13) then writes

$$\begin{aligned} u^{\text{FF}} &= \sum_{p=1}^{N_p} w_p \exp\left(-i\alpha_p\right) \frac{-i}{4}\sqrt{\frac{2}{\pi z_p}} \exp\left(i\pi/4\right) \exp\left(-iz_p\right) \\ &\approx \frac{-i}{4}\sqrt{\frac{2}{\pi z}} \exp\left(i\pi/4\right) \exp\left(-iz\right) \sum_{p=1}^{N_p} w_p \exp\left(-i\alpha_p\right) \exp\left(ik_0\hat{\boldsymbol{x}}\cdot\boldsymbol{x}_p\right). \end{aligned} \tag{4.17}$$

The interesting thing about Eq.(4.17) is that its radial (z) and azimuthal ($\hat{\boldsymbol{x}}$) dependencies are independent. The far field u^{FF} can therefore be written as

$$u^{\text{FF}} = A(z)F(\hat{\boldsymbol{x}}) \tag{4.18}$$

with the *amplitude* A defined as

$$A(z) \triangleq \frac{-i}{4}\sqrt{\frac{2}{\pi z}} \exp\left(i\pi/4\right) \exp\left(-iz\right) \tag{4.19}$$

and the *radiation pattern* F defined as

$$F(\hat{\boldsymbol{x}}) = \sum_{p=1}^{N_p} w_p \exp\left(-i\alpha_p\right) \exp\left(ik_0\hat{\boldsymbol{x}}\cdot\boldsymbol{x}_p\right). \tag{4.20}$$

At any distance z from the point sources, the relative radiation pattern in direction $\hat{\boldsymbol{x}}$ is therefore

$$\frac{|u^{\text{FF}}|}{\max_{\hat{\boldsymbol{x}}}|u^{\text{FF}}|} = \left|\frac{F(\hat{\boldsymbol{x}})}{F_{\max}}\right|. \tag{4.21}$$

The radiation pattern thus dictates in which direction $\hat{\boldsymbol{x}}$ the field has the largest amplitude. More importantly, this direction can be changed by tuning the sources weights. These conclusions are made more explicit in the following paragraph.

2-point source example Consider two point sources at the same distance $\|\boldsymbol{x}_1\| = \|\boldsymbol{x}_2\| = d$ from the origin and with the same weights $w_1 = w_2 = 1$, the radiation pattern F becomes

$$\begin{aligned} F(\hat{\boldsymbol{x}}) &= \exp\left(-i\alpha_1\right)\exp\left(ik_0\hat{\boldsymbol{x}}\cdot\boldsymbol{x}_1\right) + \exp\left(-i\alpha_2\right)\exp\left(ik_0\hat{\boldsymbol{x}}\cdot\boldsymbol{x}_2\right) \\ &= \exp\left(-i\frac{\alpha_1+\alpha_2}{2}\right)\left[\exp\left(-ik_0 d\cos\theta - i\frac{\alpha_1-\alpha_2}{2}\right) + \exp\left(ik_0 d\cos\theta - i\frac{\alpha_2-\alpha_1}{2}\right)\right] \\ &= \exp\left(-i\frac{\alpha_1+\alpha_2}{2}\right) 2\cos\left(-k_0 d\cos\theta + \frac{\Delta\alpha}{2}\right) \end{aligned} \tag{4.22}$$

where $\Delta\alpha \triangleq \alpha_2 - \alpha_1$ is the *phase difference*. The direction in which the radiated field is maximal is then given by

$$\cos\theta_M = \frac{\Delta\alpha}{2k_0 d}. \tag{4.23}$$

Imposing that the phase difference $\Delta\alpha \in]-\pi, \pi]$, then choosing $d = \lambda/4$ yields that the maximal radiating direction can be tuned between 0 and π simply by changing the phase difference. This spacing will thus be considered further in this work. The azimuthal dependency $|F(\theta)|/F_{\max}$ is represented in Figure 4.3 for different values of $\Delta\alpha$, as expected, the maximal direction of radiation changes according to Eq.(4.23).

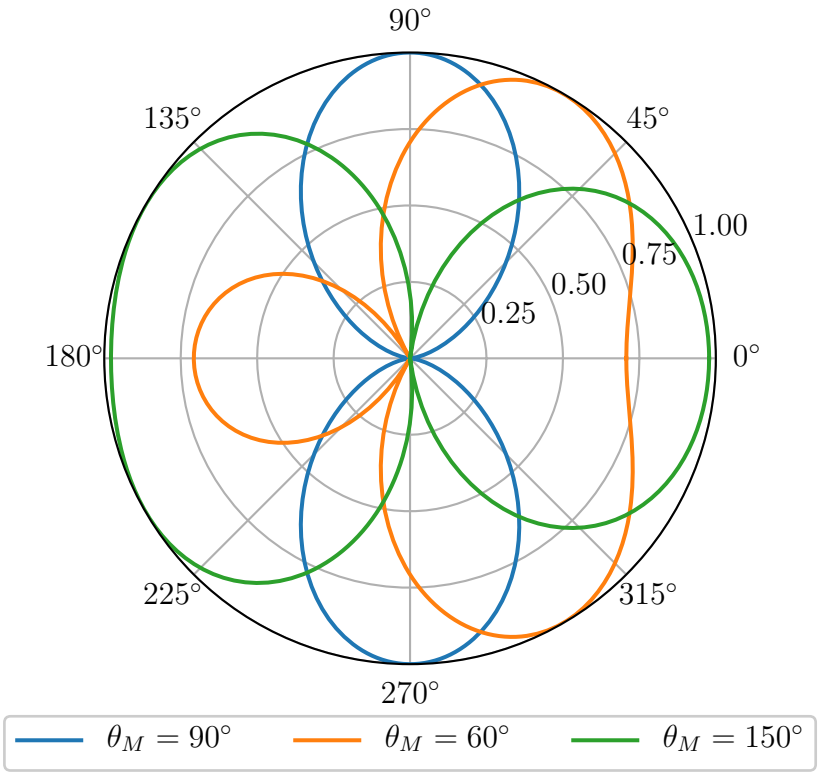


Figure 4.3: Azimuthal dependency $|F(\theta)|/F_{\max}$ for different phase difference $\Delta\alpha$. $\theta_M = 90° \rightarrow \Delta\alpha = 0$, $\theta_M = 60° \rightarrow \Delta\alpha = \pi/2$, $\theta_M = 150° \rightarrow \Delta\alpha = -\frac{\sqrt{3}}{2}\pi$.

4.3 Point source as a physical limit

Point sources are convenient here from the mathematical point of view because the analytical solution in homogeneous space is known. From a physical point of view however, such infinitesimal sources can not exist. Some physical situations are nevertheless nearly equivalent to point sources. For these particular physical sources, the elementary solutions are thus close to the non-homogeneous space Green function.
In this section, two physical approximations of point sources are discussed. These particular sources are used further in this work. The goal of this section is therefore to emphasize that, from a mathematical point of view, the phased array optimization is based on the computation of the non-homogeneous Green function, which can actually only be computed using approximations of point sources [15] [31].

Mathematically speaking, a point source is called a *Dirac distribution*. An important property of this Dirac distribution δ is that it satisfies

$$\begin{cases} \delta(\boldsymbol{x}) &= 0 \qquad \forall \boldsymbol{x} \in \mathbb{R}^2 \setminus \{0\}, \\ \displaystyle\int_{\mathbb{R}^2} \delta \, d\boldsymbol{x} &= 1. \end{cases} \tag{4.24}$$

A good approximation should thus also nearly satisfy these conditions and especially, satisfy them when some small parameter ϵ tends to zero.

Constant approximation The first approximation to be considered is the zeroth order polynomial approximation, or constant approximation

$$\delta_{\epsilon,0}(\boldsymbol{x}) = \begin{cases} \dfrac{1}{\pi\epsilon^2} & \forall \|\boldsymbol{x}\| \leq \epsilon \\ 0 & \forall \|\boldsymbol{x}\| > \epsilon. \end{cases} \tag{4.25}$$

This is the simplest approximation that satisfies Eqs.(4.24) when $\epsilon \to 0$. Physically, it simply means to replace the infinitely thin wire by a wire with a radius ϵ and a constant current density.
This approximation is referred to as CST.

Gaussian approximation The second approximation considered is the Gaussian curve

$$\delta_{\epsilon,g}(\boldsymbol{x}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\frac{\|\boldsymbol{x}\|}{\sigma}\right)^2\right) \quad \text{with } \sigma = \epsilon/4 \tag{4.26}$$

which also satisfy Eq.(4.24) when $\epsilon \to 0$. As previously, this approximation consists in considering a circular wire of radius ϵ but with an exponentially decreasing current.
This approximation is referred to as GSN.

No approximation Another possibility is to consider the source $f = \delta$, *i.e.* making no approximation. In the finite element framework described in Chapter 3, this source is then discretised in a hat function that thus depends on the mesh around the point where the Dirac is imposed. This hat function also converges to the Dirac distribution when the characteristic mesh size tends to zero. However, this approximation of the dirac distribution is not optimal, one of the reason being that it is not symmetric. Physically, this approximation consists in a mesh-dependent rectangular wire with a hat function current density, which makes no sense. This possibility is not considered in this work. More rigorous and complete treatments of these approximations can be found in [15] [31].

For both approximations, the size ϵ of the support of the function has to be determined. For these approximations to yield good results, this parameter must be such that $\epsilon \ll r_{\text{min}}$, r_{min} being the distance from the point source where the elementary solution begins to differ from the non-homogeneous space Green function. Indeed, an important result related to the Green functions for linear partial differential equations is that the solution for any source f can be written as

$$u = \int g(\boldsymbol{x}, \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y}. \tag{4.27}$$

If the source f is an approximate Dirac distribution Eq.(4.25) or Eq.(4.26) then one has in free space[5]

$$\begin{aligned} u &= \int_{\mathbb{B}(0,\epsilon)} g_0(\boldsymbol{x}, \boldsymbol{y}) \delta_\epsilon(\boldsymbol{y})\, d\boldsymbol{y} \\ &\approx g_0(\boldsymbol{x}, 0) \int_{\mathbb{B}(0,\epsilon)} \delta_\epsilon(\boldsymbol{y})\, d\boldsymbol{y} \quad (\text{if } \epsilon \ll \|\boldsymbol{x}\|) \\ &= g_0(\boldsymbol{x}, 0) \end{aligned} \tag{4.28}$$

[5]The last equality is exact for the constant approximation, nearly exact for the Gaussian approximation as a truncation is performed.

where $\mathbb{B}(0,\epsilon)$ is the centered disc of radius ϵ. The free space Green function $g_0(\boldsymbol{x},\boldsymbol{y})$ can actually be considered constant in the integral because

$$\begin{aligned} k_0\left\|\boldsymbol{x}-\boldsymbol{y}\right\| &= k_0\sqrt{\left\|\boldsymbol{x}\right\|^2+\left\|\boldsymbol{y}\right\|^2-2\boldsymbol{x}\cdot\boldsymbol{y}} \\ &= k_0\left\|\boldsymbol{x}\right\|\sqrt{1+\left(\frac{\left\|\boldsymbol{y}\right\|}{\left\|\boldsymbol{x}\right\|}\right)^2-2\left(\frac{\boldsymbol{x}\cdot\boldsymbol{y}}{\left\|\boldsymbol{x}\right\|^2}\right)} \\ &\approx k_0\left\|\boldsymbol{x}\right\|, \qquad \forall\boldsymbol{y}\in\mathbb{B}(0,\epsilon),\forall\left\|\boldsymbol{x}\right\|\gg\left\|\boldsymbol{y}\right\|. \end{aligned} \tag{4.29}$$

Consequently, the elementary solution for both constant and Gaussian sources is expected to be close to the non-homogeneous Green functions at distance $\|\boldsymbol{x}\|$ greater than their radius ϵ. Inside this radius, the elementary solution and the Green function are not necessarily equal.

Chapter 5

Optimization problems

This chapter's aim is to formulate and solve optimization problems for the source weights of phased arrays introduced in Chapter 4. Two kinds of optimization problems that can be solved using the tools described in Part I are introduced.
In Chapter 4, the weights j_p were considered to be complex-valued. To be able to formulate the optimization as real-valued problems, the real and imaginary parts of the weights are splitted as

$$j_p = j_p^0 + i j_p^1. \tag{5.1}$$

The current density $\boldsymbol{j}_a$ then expresses as

$$\boldsymbol{j}_a = \sum_{r=\{0,1\}} \sum_{p=1}^{N_p} j_p^r i^r \boldsymbol{w}_p \tag{5.2}$$

and the solution u as

$$u = \sum_{r=\{0,1\}} \sum_{p=1}^{N_p} j_p^r i^r u_p. \tag{5.3}$$

For the sake of simplicity, in the following sections the source weights $j_p^r \; \forall r \in \{0,1\}, \; \forall p \in [1, N_p]$ are gathered into vectors denoted by $\underline{j}$ and defined as

$$\underline{j}^T = \left[j_1^0, \, j_2^0, \, \ldots, \, j_{N_p-1}^0, \, j_{N_p}^0, \, j_1^1, \, j_2^1, \, \ldots, \, j_{N_p-1}^1, \, j_{N_p}^1 \right]. \tag{5.4}$$

5.1 Problem based on power conservation

Power conservation is expressed through the real part of Poynting's theorem, which can be obtained from the frequency domain Maxwell's equations and takes the following form:

$$P_s = L + R \tag{5.5}$$

where P_s is the *injected power* inside a control volume V, L are the *losses* inside this same volume and R is the *radiated power* through the surface S of this control volume. All these terms are expressed in [W] for three dimensional fields but in [W/m] for two dimensional cases. Proofs of this theorem for both three and two dimensional cases and the expressions of P_s, L and R are given in Appendix B.7.

Using the particular source configuration of arrays Eq.(5.2), the different terms of Poynting's theorem can be written in matrix form.
Indeed, the power injected by the source is[1]

$$\begin{aligned}
P_s &\triangleq -\int_V \text{Re}\left(u {j_a}^*\right)\, dV \\
&= -\sum_{r=\{0,1\}}\sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}}\sum_{q=1}^{N_p} j_q^s \int_V \text{Re}\left(i^r u_p (-i)^s {w_q}^*\right)\, dV \\
&= \sum_{r=\{0,1\}}\sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}}\sum_{q=1}^{N_p} j_q^s \left[-\text{Re}\left(i^r(-i)^s \int_V u_p {w_q}^*\, dV\right)\right] \\
&\triangleq \sum_{r=\{0,1\}}\sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}}\sum_{q=1}^{N_p} j_q^s \left[-\text{Re}\left(i^r(-i)^s \left[\underline{\underline{P_s^b}}\right]_{pq}\right)\right] \qquad (5.6)\\
&\triangleq \underline{j}^T \underline{\underline{P_s}}\, \underline{j}, \qquad (5.7)
\end{aligned}$$

the total radiated power is

$$\begin{aligned}
R &\triangleq \frac{1}{\omega\mu_0}\int_S \text{Im}\left(u\mathbf{grad}\, u^*\right)\cdot\hat{\boldsymbol{n}}\, dS \\
&= \sum_{r=\{0,1\}}\sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}}\sum_{q=1}^{N_p} j_q^s \frac{1}{\omega\mu_0}\int_S \text{Im}\left(i^r(-i)^s u_p \mathbf{grad}\, {u_q}^*\right)\cdot\hat{\boldsymbol{n}}\, dS \\
&= \sum_{r=\{0,1\}}\sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}}\sum_{q=1}^{N_p} j_q^s \left[\frac{1}{\omega\mu_0}\text{Im}\left(i^r(-i)^s \int_S u_p \mathbf{grad}\, {u_q}^*\cdot\hat{\boldsymbol{n}}\, dS\right)\right] \\
&\triangleq \sum_{r=\{0,1\}}\sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}}\sum_{q=1}^{N_p} j_q^s \left[\frac{1}{\omega\mu_0}\text{Im}\left(i^r(-i)^s \left[\underline{\underline{R^b}}(S)\right]_{pq}\right)\right] \qquad (5.9)\\
&\triangleq \underline{j}^T \underline{\underline{R}}(S)\, \underline{j}, \qquad (5.10)
\end{aligned}$$

and the total losses are

$$\begin{aligned}
L &\triangleq -\omega\int_V \text{Im}\left(\tilde{\epsilon}\right)\|u\|^2\, dV \\
&= -\omega\int_V \text{Im}\left(\tilde{\epsilon} u u^*\right)\, dV \\
&= \sum_{r=\{0,1\}}\sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}}\sum_{q=1}^{N_p} j_q^s \left[-\omega\int_V \text{Im}\left(i^r(-i)^s\left(\tilde{\epsilon} u_p {u_q}^*\right)\right)\, dV\right]
\end{aligned}$$

[1]The submatrix

$$\left[\underline{\underline{P_s}}\right]_{pq} \triangleq \int_V u_p w_q\, dV \qquad (5.8)$$

is actually symmetric because of the Rayleigh-Carson reciprocity theorem [27] (if ${w_p}^* = \pm w_p \forall p \in [1, N_p]$). Intuitively, as far as antennas are concerned, this theorem states that emission (when source is p and receiver is q) and reception (when source is q and receiver is p) involve the same exchange of power.

$$= \sum_{r=\{0,1\}} \sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}} \sum_{q=1}^{N_p} j_q^s \left[-\omega \mathrm{Im} \left(i^r (-i)^s \int_V \tilde{\epsilon} u_p {u_q}^* \, dV \right) \right]$$

$$\triangleq \sum_{r=\{0,1\}} \sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}} \sum_{q=1}^{N_p} j_q^s \left[-\omega \mathrm{Im} \left(i^r (-i)^s \left[\underline{\underline{L}}^b(V) \right]_{pq} \right) \right] \tag{5.11}$$

$$\triangleq \underline{j}^T \underline{\underline{L}}(V) \, \underline{j}. \tag{5.12}$$

If the control volume V on which the theorem is applied contains all the sources, then all the matrix are (semi-)positive definite. Indeed, by definition, the loss matrix cannot be negative, then if there exists a weight vector $\underline{j}$ such that $P_s < 0$, it means that the radiated power R is also negative. A negative radiated power means that power comes from outside the volume towards the sources, which is not physical because there are no source outside the volume.
The three matrices introduced previously are not symmetric. However, because they appear in quadratic forms, only their symmetric parts can be considered. Indeed, for any matrix $\underline{\underline{A}}$ and vector $\underline{x}$,

$$\underline{x}^T \underline{\underline{A}} \, \underline{x} = \underline{x}^T \left(\frac{\underline{\underline{A}} + \underline{\underline{A}}^T}{2} + \frac{\underline{\underline{A}} - \underline{\underline{A}}^T}{2} \right) \underline{x} = \underline{x}^T \left(\frac{\underline{\underline{A}} + \underline{\underline{A}}^T}{2} \right) \underline{x}. \tag{5.13}$$

Consequently, in the following developments, only the symmetric part of the power related matrices are considered.

Matrix power conservation In matrix form, the power conservation then writes

$$\underline{j}^T \left(\underline{\underline{P_s}} - \underline{\underline{R}}(S) - \underline{\underline{L}}(V) \right) \underline{j} = 0. \tag{5.14}$$

This statement has of course to be verified for all current distributions, hence for all $\underline{j}$. The only possibility is then that $\underline{\underline{P_s}} - \underline{\underline{R}} - \underline{\underline{L}}$ is skew symmetric.

Global phase invariance Consider u^* the exact solution to the Helmholtz equation with Sommerfeld radiation condition for a distribution of sources $\boldsymbol{j}_a$. By linearity the couple

$$\{u^* P \exp(-i\alpha), \boldsymbol{j}_a P \exp(-i\alpha)\} \tag{5.15}$$

is also a solution for any $P > 0$ and any $\alpha \in]-\pi, \pi]$. By definition the latter couples yield the same injected power P_s, radiated power R and losses L as the original solution if $P = 1$. Therefore if $\boldsymbol{j}_a$ is found to be an optimal distribution of sources then automatically $\boldsymbol{j}_a \exp(-i\alpha)$ is also optimal. It has thus be chosen to fix a phase reference by imposing

$$j_{N_p}^1 = 0. \tag{5.16}$$

The weight vectors have therefore the size $2N_p - 1$ instead of $2N_p$.

Using these quantities or related ones, interesting optimization problems can easily be formulated. For example, it could be interesting to maximize the power radiated through some surface S_r, while keeping the injected power constant, the problem to solve is then

$$\boxed{\max_{\underline{j} \in \mathbb{R}^{2N_p - 1}} R(S_r) \quad \text{with } P_s = P_0.} \tag{5.17}$$

Similarly in some situations, the dissipated power (*i.e.* the losses) in a volume V_l must be limited. The problem then becomes

$$\boxed{\max_{\underline{j}\in\mathbb{R}^{2N_p-1}} R(S_r) \quad \text{with } P_s \leq P_0 \quad \text{and } L(V_l) \leq L_0.} \tag{5.18}$$

This more general problem is however equivalent to the previous one in a important particular case. Consider $\mathbb{R}_{L_0}^{2N_p-1}$ the subset of all losses-compatible $\left(\mathbb{R}_{L_0}^{2N_p-1} = \{\underline{j} \in \mathbb{R}^{2N_p-1} | L \leq L_0\}\right)$ vectors $\underline{j}$ (resp. $\mathbb{R}_{P_0}^{2N_p-1}$ the subset of all power-compatible $\left(\mathbb{R}_{P_0}^{2N_p-1} = \{\underline{j} \in \mathbb{R}^{2N_p-1} | P \leq P_0\}\right)$ vectors $\underline{j}$), if $\forall \underline{j} \in \mathbb{R}_{L_0}^{2N_p-1}$ (resp. $\mathbb{R}_{P_0}^{2N_p-1}$) $P_s \ll P_0$ (resp. $L \ll L_0$) then the constraint on the injected power (resp. the losses) is invisible and the constraint on the losses (resp. injected power) becomes an equality constraint. The problem is then equivalent to the previous one. Using the three quantity P_s, $L(V_l)$, $R(S_r)$, other power budget optimization can easily be formulated.

5.2 Problems based on equivalent arrays

When doing electromagnetic compatibility studies, it is not rare that several complex electromagnetic components are brought together in a restricted area. The complexity of the whole system is therefore very high and it can be useful to substitute the complex constituents by equivalent arrays that have the same radiating properties in the far field. To construct this model, one idea is to look at the difference between the real radiation pattern and the radiation pattern of the equivalent array. This difference can be for example quantified by a single indicator to be the L_2 norm of the difference between the real and the equivalent field, *i.e.*

$$D \triangleq \int_S \|u - u_S\|^2 \, dS. \tag{5.19}$$

Once again using the superposition principle for the electric field Eq.(4.4) yields

$$\begin{aligned}
D &\triangleq \int_S \|u - u_S\|^2 \, dS \\
&= \int_S \|u\|^2 \, dS - 2\text{Re}\left(\int_S u u_S{}^* \, dS\right) + \int_S \|u_S\|^2 \, dS \\
&= \sum_{r=\{0,1\}} \sum_{p=1}^{N_p} j_p^r \sum_{s=\{0,1\}} \sum_{q=1}^{N_p} j_q^s \text{Re}\left[i^r (-i)^s \int_S u_p u_q{}^*\right] \\
&\quad - 2 \sum_{r=\{0,1\}} \sum_{p=1}^{N_p} j_p^r \text{Re}\left[i^r \int_S u_p u_S{}^* \, dS\right] + \int_S \|u_S\|^2 \, dS \\
&\triangleq \underline{j}^T U \, \underline{j} - 2\underline{j}^T \underline{s} + c.
\end{aligned} \tag{5.20}$$

The related optimization problem is then

$$\boxed{\min_{\underline{j}\in\mathbb{R}^{2N_p-1}} D.} \tag{5.21}$$

5.3 Solution techniques

In this section, solution techniques for problems of the forms Eq.(5.17) or Eq.(5.21) are derived. Problems of the form Eq.(5.18) are not considered anymore in this work as their solution is not straightforward and because they can be simplified using physically based arguments.

5.3.1 Pure quadratic equality constrained problems

Consider the optimization problem

$$\boxed{\max_{\underline{j}\in\mathbb{R}^{2N_p-1}} \underline{j}^T \underline{\underline{A}}\, \underline{j} \quad \text{with } \underline{j}^T \underline{\underline{C}}\, \underline{j} = 1} \tag{5.22}$$

where $\underline{\underline{A}}$ and $\underline{\underline{C}}$ are positive definite symmetric matrices. Using Lagrange multiplier technique, the problem Eqs.(5.22) is equivalent to solve

$$\max_{\underline{j}\in\mathbb{R}^{2N_p-1},\, \lambda\in\mathbb{R}} \underline{j}^T \underline{\underline{A}}\, \underline{j} - \lambda \left(\underline{j}^T \underline{\underline{C}}\, \underline{j} - 1 \right). \tag{5.23}$$

The extremum value of this function cancels the gradient w.r.t. $\underline{j}$ which yields

$$\underline{\underline{A}}\, \underline{f} = \lambda \underline{\underline{C}}\, \underline{f}. \tag{5.24}$$

The solution to the generalized eigenvalue problem Eq.(5.24) are denoted by $\{\lambda_n, \underline{j}_n\}$. The cancellation of the gradient of the cost function w.r.t. λ then gives the metric to normalize the eigenvectors. Indeed, one has

$$\underline{j_n}^T \underline{\underline{C}}\, \underline{j_n} = 1. \tag{5.25}$$

Finally, by inserting these results into Eq.(5.23), the problem reduces to

$$\max_{\underline{j}\in\mathbb{R}^{2N_p-1},\, \lambda\in\mathbb{R}} \underline{j}^T \underline{\underline{A}}\, \underline{j} - \lambda \left(\underline{j}^T \underline{\underline{C}}\, \underline{j} - 1 \right) = \max_{\underline{j}\in\mathbb{R}^{2N_p-1},\, \lambda\in\mathbb{R}} \lambda \underline{j}^T \underline{\underline{C}}\, \underline{j} = \max_{\lambda\in\mathbb{R}} \lambda. \tag{5.26}$$

such that the solution to Eqs.(5.22) is the *principal eigenvector* of $\underline{\underline{C}}^{-1} \underline{\underline{A}}$ normalized by $\underline{\underline{C}}$. The fact that the matrices are symmetric positive definite guarantees that all the eigenvalues are real-valued and positive. Developments are identical if $\max$ is substituted by $\min$ but the smallest eigenvalue should then be kept.

5.3.2 Quadratic unconstrained problems

Consider the optimization problem

$$\min_{\underline{j}\in\mathbb{R}^{2N_p}} \underline{j}^T \underline{\underline{A}}\, \underline{j} - 2\underline{j}^T \underline{b} + c \tag{5.27}$$

where $\underline{\underline{A}}$ is again a symmetric positive definite matrix, $\underline{b}$ is a vector and c a constant. The minimum of the cost function cancels the gradient w.r.t. $\underline{j}$, *i.e.*

$$2\underline{\underline{A}}\, \underline{j} - 2\underline{b} = 0. \tag{5.28}$$

The solution is then straightforward and is

$$\underline{j_m} = \underline{\underline{A}}^{-1} \underline{b}. \tag{5.29}$$

Part III

Applications

Chapter 6

Numerical validation of elementary solution

Part I introduces methods to solve, among others, equations similar to Eqs(4.5) whose solutions are called elementary solutions. The focus of this part is to illustrate the performance of these methods on simple model problems for which the exact solution is known. In particular, the effect of absorbing boundary conditions, point source approximation, transmission conditions and the iterative solver are investigated in this chapter. The quantities of interest to measure the difference between a numerical solution u and a reference solution u^* are the relative error e defined as

$$e = \left| \frac{u - u^*}{u^*} \right| \tag{6.1}$$

and its L_2 and L_∞ norms, *i.e.*

$$L_2(e) = \sqrt{\frac{1}{A_\Omega} \int_\Omega e^2 d\Omega} \quad \text{and} \quad L_\infty(e) = \max_{\boldsymbol{x} \in \Omega} e \tag{6.2}$$

where A_Ω is the area of Ω.

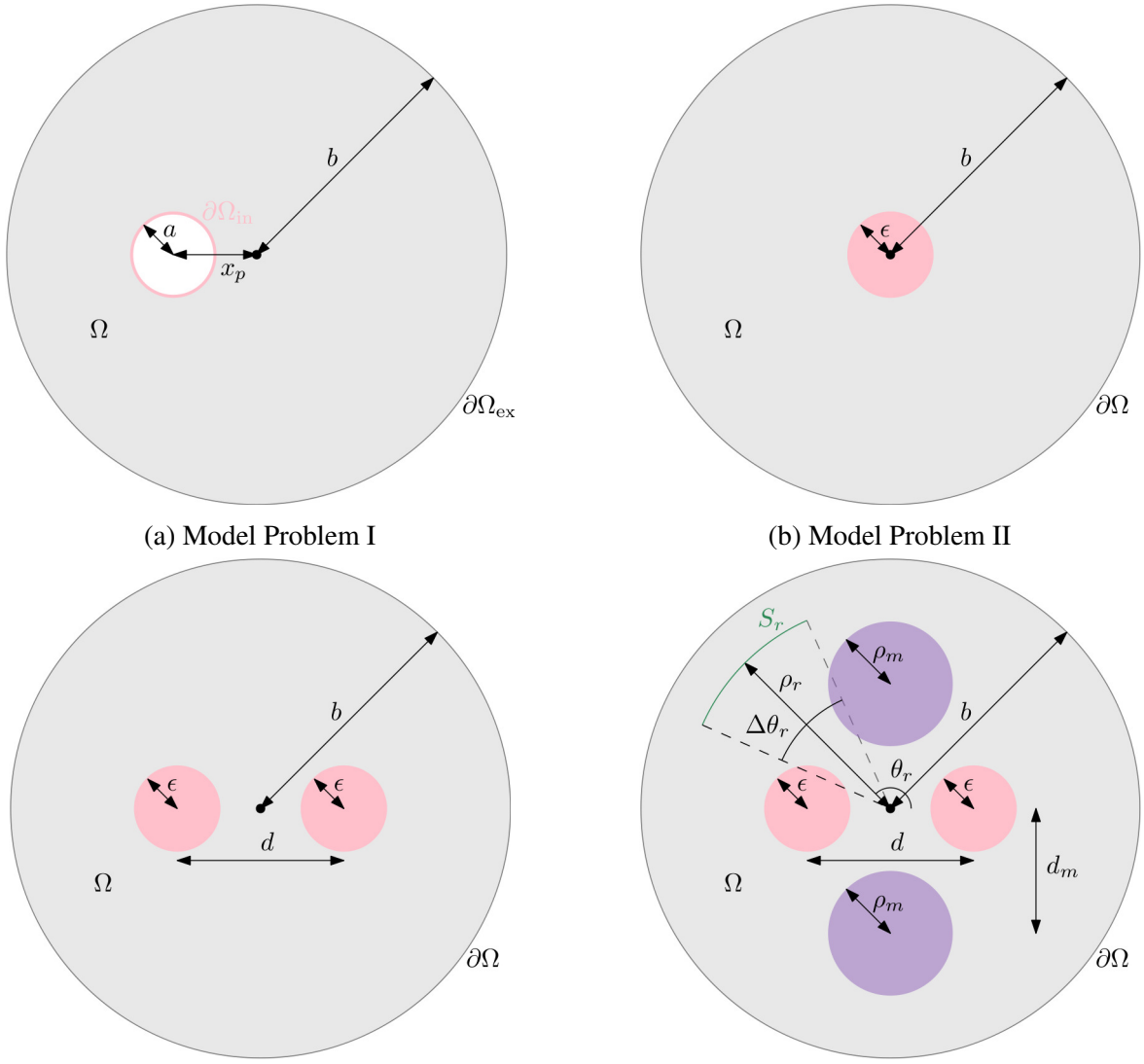
6.1 Absorbing boundary conditions

As emphasized in Section 1.6, the elementary problem Eqs(4.5) cannot be discretized as such but the Sommerfeld radiation condition must be substituted by an approximate condition at a finite distance. In this section, the effect of this approximation is briefly studied.

Model Problem I The model problem used here is a single point source, *i.e.* a 1-source array, in free space. The latter problem is however not suited to study the performance of the absorbing boundary conditions alone because a second approximation is required to model the point source. To get rid of this second approximation, the region around this point source is withdrawn from the domain and a Dirichlet boundary condition is imposed on the interior surface. The problem solved in this section is then

$$\begin{cases} \left(-\operatorname{div} \mathbf{grad} \; - k_0^2\right) u = 0 & \forall \boldsymbol{x} \in \Omega, \\ \hat{\boldsymbol{n}} \cdot \mathbf{grad}\, u + \mathcal{B} u = 0 & \forall \boldsymbol{x} \in \partial\Omega_{\text{ex}}, \\ u = u^*(\boldsymbol{x}) & \forall \boldsymbol{x} \in \partial\Omega_{\text{in}} \end{cases} \tag{6.3}$$

where $\partial\Omega_{\text{in}}$ and $\partial\Omega_{\text{ex}}$ are two circles of radii a and b and Ω is the region between them. Obviously, the exact solution to the unbounded problem (with the Sommerfeld radiation condition) is the Green function $u^* = g_0$. For the numerical experiments, the wave number k_0 is set to 1 such that the wavelength $\lambda = 2\pi$. The characteristic element size h is taken to $\lambda/50$ such that errors due to the discretisation are small. The geometry of this model problem is given in Figure 6.1a while the numerical parameters are summarized in Table C.1 (*cfr* Appendix).



(a) Model Problem I

(b) Model Problem II

(c) Model Problem III

(d) Model Problem IV

Figure 6.1: Geometries relative to model problems.

Results The map of the error and its norms are given in Figure 6.2 and in Table 6.1 for different configurations. The absorbing boundary conditions is located at either $b = 6 \approx \lambda$ or $b = 30 \approx 5\lambda$. The point source is considered either at the center of the domain or off-center by $x_p = -3 \approx -\lambda/2$. The inner Dirichlet boundary has a radius $a = 0.6 = \lambda/10$.

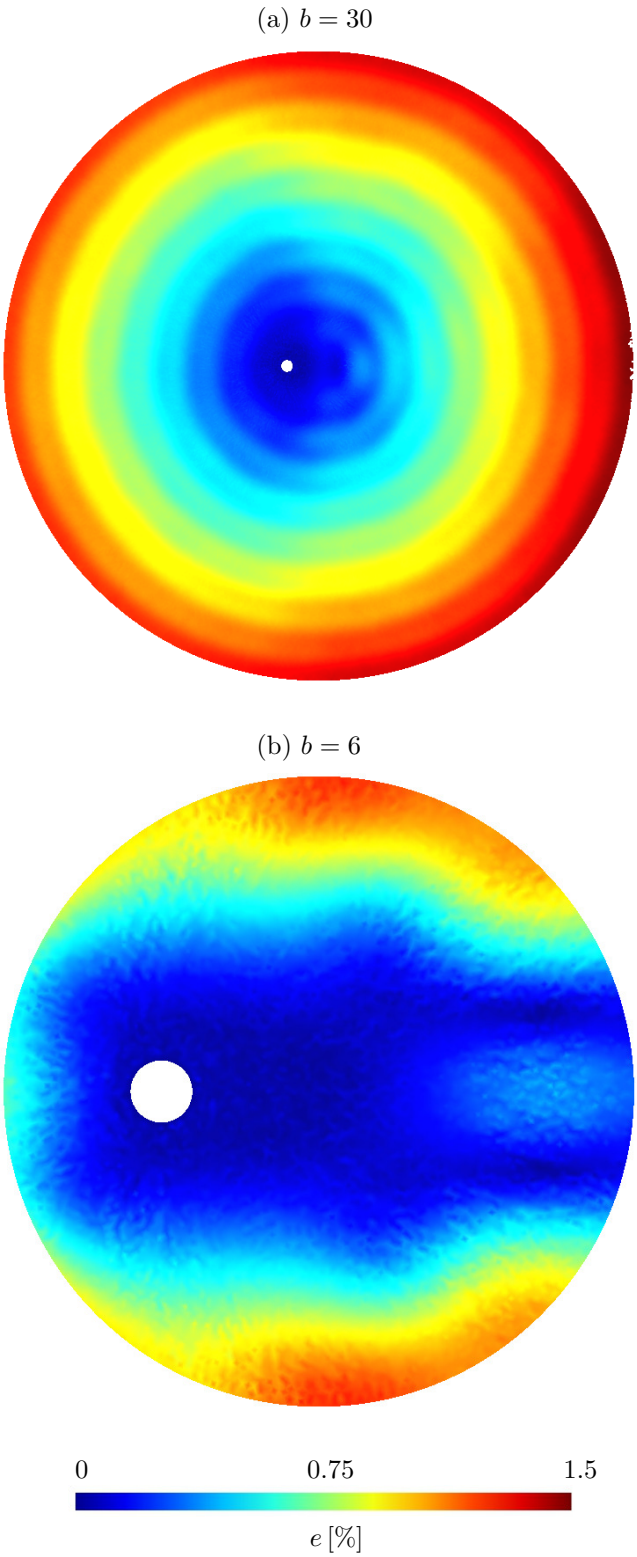


Figure 6.2: Relative error e for two position of the ABC. The point source is off-center by $x_p = -3$. Second order Taylor absorbing boundary condition are used.

	$\{x_p, y_p\} = \{0, 0\}$		$\{x_p, y_p\} = \{-3, 0\}$	
	$b = 6$	$b = 30$	$b = 6$	$b = 30$
CT0	0.16(0.3)	0.9(1.3)	7.7(17.2)	1.0(1.8)
CT2	0.22(0.3)	0.9(1.3)	0.6(1.2)	0.9(1.5)

Table 6.1: $L_2(L_\infty)$ norm of the relative error e for several configurations. Values are expressed in %.

Discussion When the source is centered, the solution is symmetric and thus the only difference between both Taylor approximations is terms proportional to the curvature squared because the tangential derivative vanishes, hence the very small difference. When the source is off-center, the solution is no more constant on the circular boundary. A non negligible difference then appears between both Taylor approximations. This difference however diminishes for $b = 30$ and the reason for that is the same than when the source is centered. Indeed, as it is shown graphically in Figure 6.3, the further the boundary, the smaller the angle between the wavefront and the boundary, and consequently the closer the solution to the centered case.

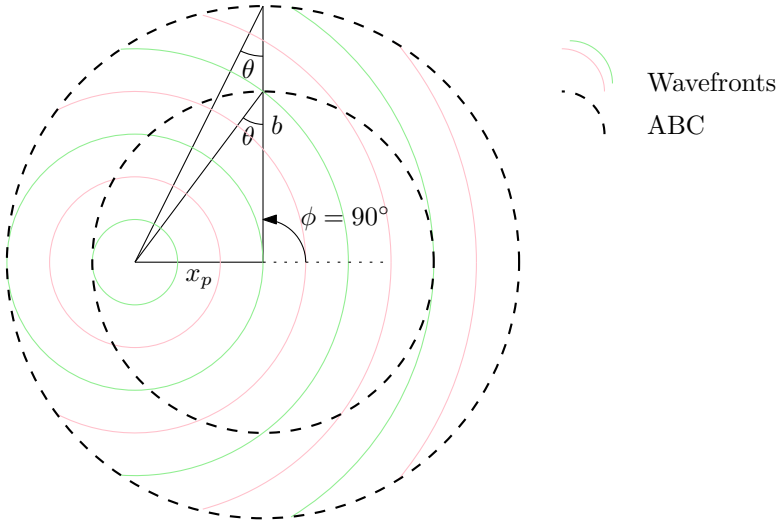


Figure 6.3: Graphical representation of the wavefronts of the far field Green function.

From Table 6.1 it could be wrongly concluded that there is no interest to set the boundary condition further from the source as both L_2 and L_∞ are not strongly affected when CT2 is used. However, when looking at the error map in Figure 6.2, it can be seen that the error is concentrated near the boundary. Thus if this boundary is far from the region of interest, *i.e.* around the source in this case, then the error locally decreases. It is also interesting to emphasize that the incidence angle θ (*i.e.* the angle formed by the normal to the boundary and the circular wavefront) decreases as $\frac{x_p}{b}$ decreases, for example, when $\phi = 90°$ then $\theta = \tan^{-1}\left(\frac{x_p}{b}\right)$ thus either $\theta \approx 6°$ or $\theta \approx 27°$[1]. In Section 1.6, it

[1]Actually the maximum incidence angle θ_{max} occurs at $\phi_{\text{max}} = \cos^{-1}\left(-\frac{x_p}{b}\right)$ and is given by

$$\theta_{\text{max}} = \frac{\frac{x_p}{b}}{\sqrt{1 - \left(\frac{x_p}{b}\right)^2}}. \tag{6.4}$$

is shown that the reflection (and thus the error) increases with the incidence angle θ. Results from Section 1.6 can however not be applied readily here as a spherical wave is a superposition of plane waves. As showed graphically in Figure 6.3 the ratio $\frac{x_p}{b}$ explains why the error is nearly symmetric for $b = 30$ and not for $b = 6$. Indeed, for $b = 30$ the incidence angle is nearly the same in all directions while this is not the case for $b = 6$.
As a conclusion, the external boundary should be placed as far as possible and CT2 should be used to have the most accurate results.

6.2 Point source approximation

In Section 4.3, the link between theoretical point source and two physical sources is introduced. In this section, the correspondence between free space Green function and elementary solution is studied.

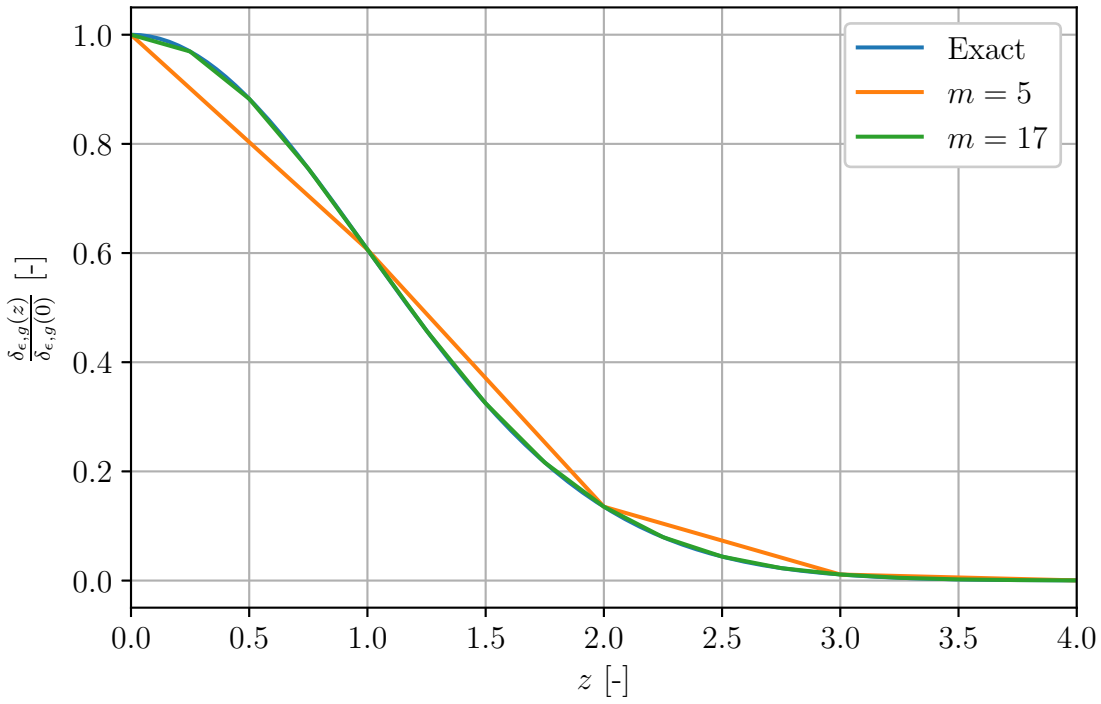


Figure 6.4: Gaussian curve approximation with $m = 5$ and $m = 17$ points.

Model Problem II The model problem of this section is similar to the preceding one (*i.e.* Model Problem I), the only two differences are that the Dirichlet condition is replaced by a physical source and that this source is always centered. For the sake of simplicity, the constant $-i\omega\mu_0 j_0 A_\delta$ is set to 1. For the numerical simulations, the characteristic mesh size is $h = \lambda/50$ near the absorbing boundary condition but $h_s = \frac{\epsilon}{m}$ inside the point source approximation. Indeed, as the characteristic mesh size should be small w.r.t. the smallest geometric characteristic length, it is therefore risky to take $m = 1$. This new parameter m is actually important, especially in the case of the Gaussian

approximation. By taking a mesh size too large, the Gaussian curve will not be discretised correctly as can be seen for $m = 5$ and $m = 17$ in Figure 6.4. The geometry of this model problem is given in Figure 6.1b while the numerical parameters are summarized in Table C.1 (*cfr* Appendix).

Results As previously, L_2 and L_∞ norms of the error e are given in Table 6.2 for both physical point source approximations, for $m = 5$ and $m = 17$ and for a source radius $\epsilon = 0.6 \approx \lambda/10$ and $\epsilon = 0.3 \approx \lambda/20$. The map of the error e is given in Figure 6.5.

	$\epsilon = 0.6 \approx \lambda/10$		$\epsilon = 0.3 \approx \lambda/20$	
	$m = 5$	$m = 17$	$m = 5$	$m = 17$
CST	5.25(96.7)	4.62(96.7)	1.85(94.7)	1.26(94.5)
GSN	1.39(94.2)	1.24(94)	0.45(92.2)	0.39(92)

Table 6.2: $L_2(L_\infty)$ norm of the relative error for several configurations. Values are expressed in %.

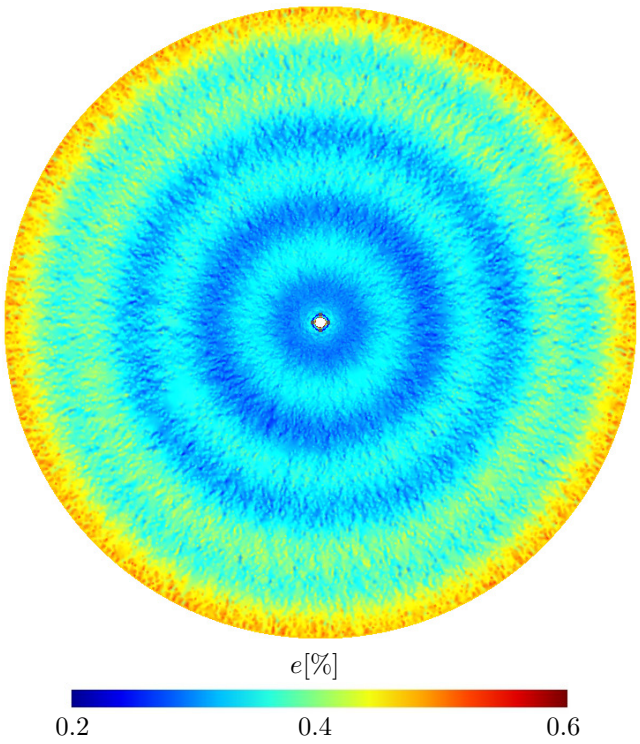


Figure 6.5: Relative error e for Model Problem II.GSN source with $m = 5$ and $\epsilon = 0.3 \approx \lambda/20$ are used. The white region inside the source is off scale, the maximal value inside this white region is 92.2%.

Discussion From Table 6.2, it appears that the maximal error is huge and does not decreases with m or ϵ. However this error always appears inside the point source approximation (*cfr* Figure 6.5) which makes sense as the exact and approximate problems strongly differ in that region. This maximal error is actually expected from Section 4.3. Moreover, as can be seen in Table 6.2 and Figure

6.5, the error is small and uniform in the region outside the source radius.
The number of points per source radius m does not seem to strongly influence the error. This fact could be surprising at first sight because of the bad Gaussian approximation that $m = 5$ yields (*cfr* Figure 6.4). The results are actually still good because the approximation of the Gaussian curve (*i.e.* the orange curve in Figure 6.4) is also an approximation of a point source.
The source radius allows to efficiently reduce the difference between the free space Green function and the elementary solution. For the Gaussian approximation, with $\epsilon = 0.3 \approx \lambda/20$ the errors are of the same order than those obtained with the absorbing boundary conditions only, it is therefore useless to decrease this radius further.

As a conclusion, the important parameter, as far as convergence towards Green function is concerned, is the size of the support for both approximations. The m parameter is only important to represent geometrically physical circular wires. Within the Dirac approximations (*i.e.* $\|\boldsymbol{x}\| \leq \epsilon$), the solution is however far from the Green function. The approximation used can therefore influence any quantities related to the fields very close to the sources, *e.g.* the injected power P_s.

6.3 Transmission condition and iterative solver

Now that approximations independent of domain decomposition methods have been discussed, the performance of the different transmission conditions and of both iterative solvers are investigated. First the transmission conditions in free space are considered. Then the transmission condition for complex wavenumber are used on a realistic example.

6.3.1 Real wavenumber

Model Problem III The model problem for real wavenumber consists in a 2-source array. Sources are separated by $d = \lambda/2 = \pi$ and have unitary weights. The domain is divided into $N_d = 5$ subdomains. The geometry of this model problem is given in Figure 6.1c while the numerical parameters are summarized in Table C.1 (*cfr* Appendix).

Results The relative residual $fracr_n r_1$ as a function of the number of iterations n is given in Figure 6.6, for both Jacobi and Krylov solvers and for T0BC, T2BC and OO2BC transmission conditions. The L_2 norm of the error e as a function of the tolerance ϵ on the relative residual for one of the subdomains is given in Figure 6.7 for Krylov solver and OO2 transmision condition. The whole map of the error e in the same configuration is given in Figure C.1 (*cfr* Appendix). Finally, the solution is given in Figure C.2 (*cfr* Appendix).

Discussion From Figure 6.6, it can be concluded that OO2BC transmission conditions are the most efficient for both solvers. Both T0BC and T2BC transmission conditions converge when used with Krylov solver but do not converge for the Jacobi solver. This observation can be explained by the convergence rates obtained for the two-subdomain problem in Section 2.4. Indeed, it was found that modes such that $k_y \geq k_0$ diverge. It is important to highlight that the convergence rates found in Section 2.4 are strictly valid only for the two-subdomain problem and for Jacobi scheme. That is why, using Krylov solver, convergence is still reachable for T0BC and T2BC transmission conditions.

As the quantity of interest is the volume unknown u, it is necessary to link the tolerance on the residuals ϵ_{it} (which is computed on surface unknowns) to an estimate of the error on the volume unknown u. That is precisely what is done in Figure 6.7. As can be seen in that figure, when the tolerance ϵ_{it} is smaller than 10^{-3}, the error seems not to decrease anymore. This lower bound on the error is actually of the order of the error found previously in Section 6.1 or 6.2, because of absorbing boundary conditions or point source approximation. To be sure that there is not peak of error nowhere in one of the $N_d = 5$ subdomains, a map of the error e is given in Figure C.1. As expected, there is no error peak except inside the sources.
Finally the solution is given in Figure C.2. The radiation pattern that is predicted in Section 4.2 is easily identified.

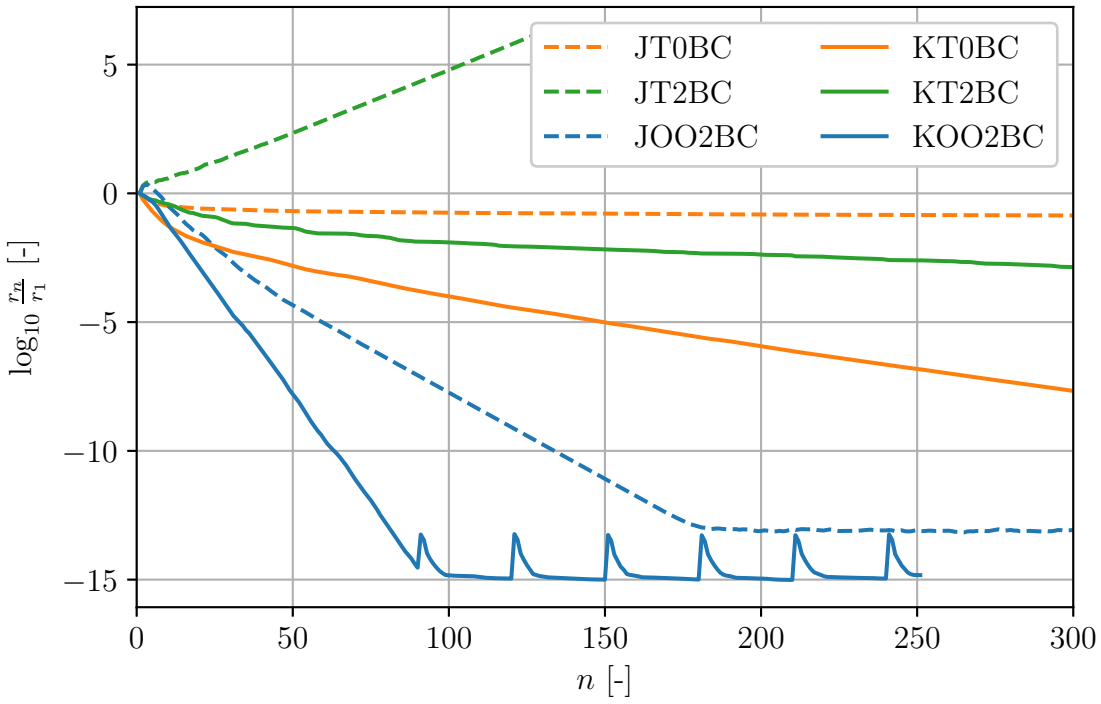


Figure 6.6: Relative residual $\frac{r_n}{r_1}$ as a function of the number of iterations n for both Jacobi (J-) and Krylov (K-) solvers and for T0BC, T2BC and OO2BC transmission conditions.

6.3.2 Complex wavenumber

The problem considered here is actually the realistic application that is considered in Chapter 8 for an elementary source. A complete description of the model is provided in that chapter. For now, it is only interesting to know that the model is composed of five different media which are not magnetic but which have different permittivities that depend on the frequency. These media are surrounded by empty space.
As it is emphasized in Section 2.4, the complex nature of the wavenumber has to be considered only when the equivalent conductivity $\tilde{\sigma}$ is non negligible. To be able to see the influence of this equivalent conductivity, two frequencies are used; $f = 100$ [MHz] and $f = 3$ [GHz]. At $f = 100$ [MHz], the imaginary and the real part of the permittivity are more or less equal while at $f = 3$ [GHz], the

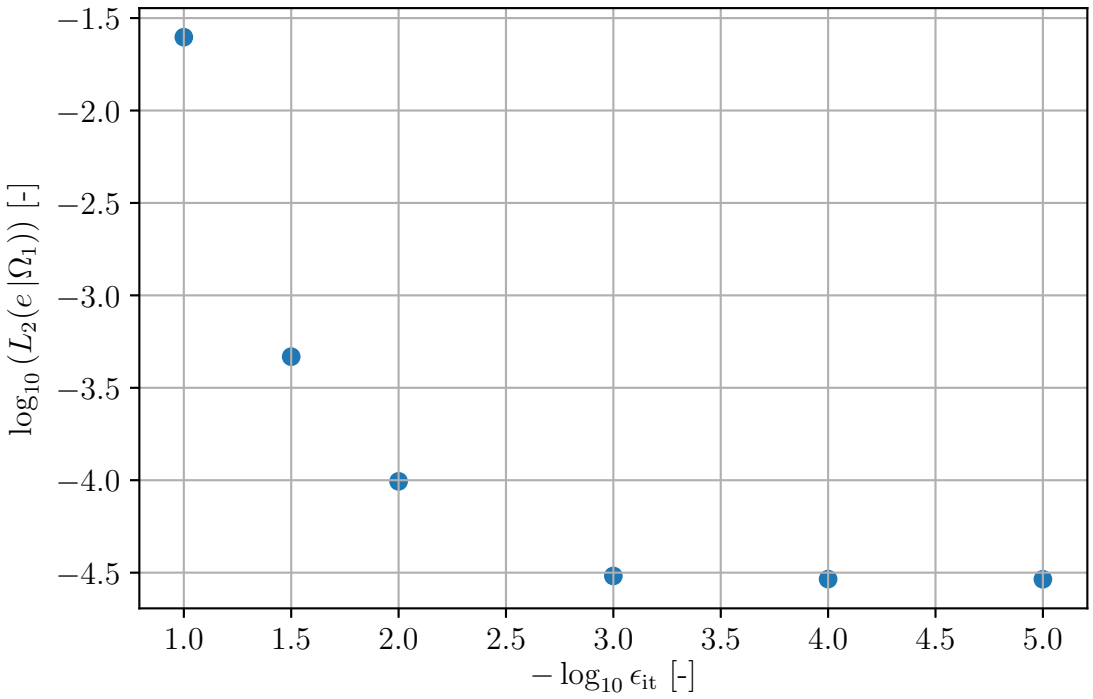
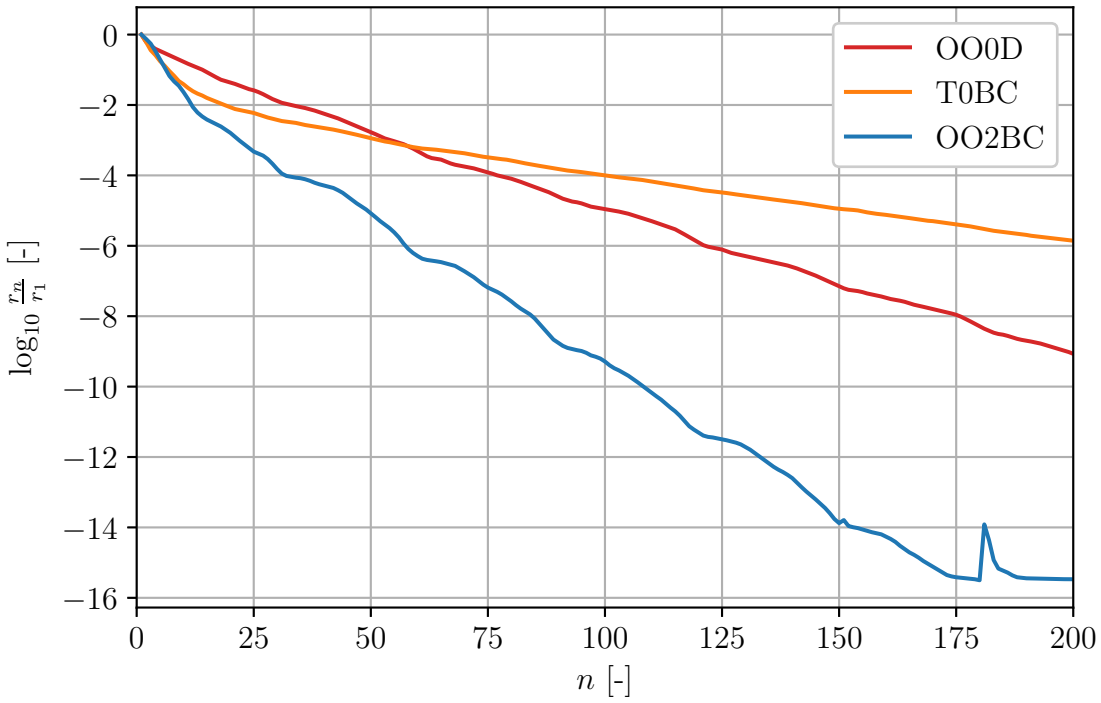


Figure 6.7: L_2 norm of the error e as a function of the tolerance ϵ_{it} on the relative residual for one of the subdomains and for Krylov solver and OO2 transmission conditions.

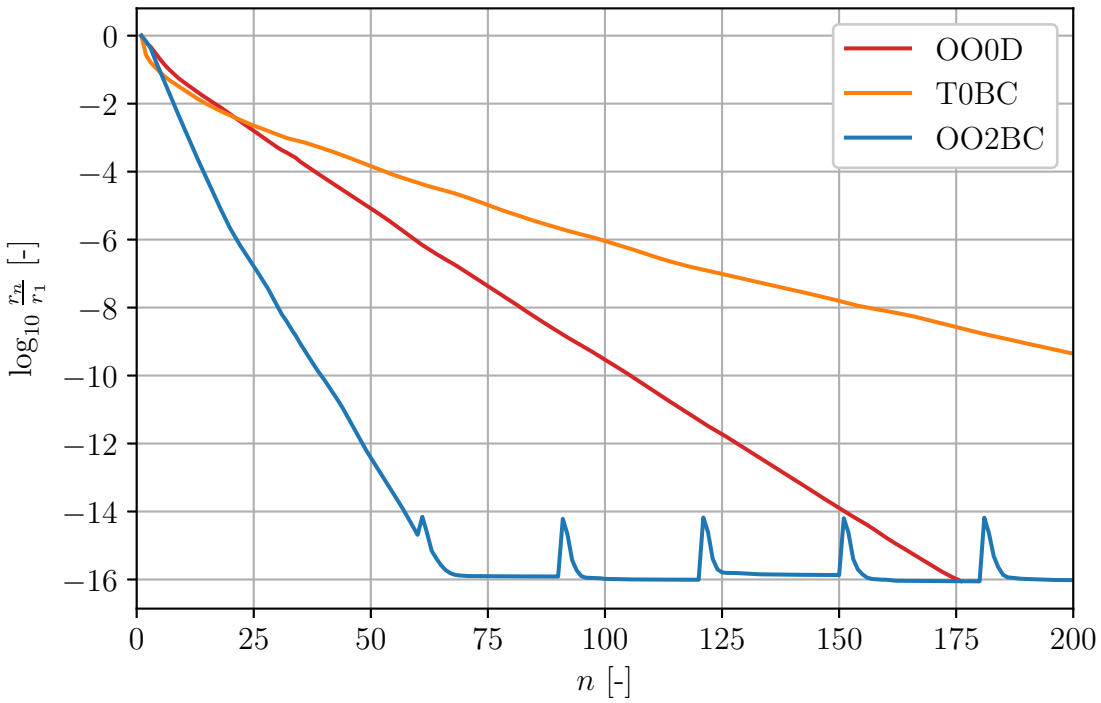
imaginary part is slightly smaller (up to 5 times smaller). The exact value of permittivities and the ratio $\frac{\tilde{\sigma}}{\omega \text{Re}(\tilde{\epsilon})}$ are given for each media in Figure 8.3, C.4 and C.5 (*cfr* Appendix).
For the numerical simulations, the characteristic mesh sizes are taken such that there are 60 (resp. 120) points per wavelength[2] at $f = 100$ [MHz] in empty space (resp. in the media) and 12 points per wavelength[2] at $f = 3$ [GHz] (both in empty space and in the media). The current carried by the physical source is $j_0 A_\delta = 1$ [A]. The geometry of this model problem is given in Figure 8.1 while the numerical parameters are summarized in Table C.1 (*cfr* Appendix).

Results The residual as a function of the iteration number is given in Figure 6.8a and 6.8b for both frequencies. T0BC, OO2BC and OO0D transmission conditions are used inside the conductive media while OO2BC is always used in the surrounding empty space. Krylov solver only is used.

[2]A reference wavelength is defined from Figure 8.2 to be $\lambda_m = 0.3$ [m] at $f = 100$ [MHz] and $\lambda_m = 0.01$ [m] at $f = 3$ [GHz].



(a) $f = 100$ [MHz]



(b) $f = 3$ [GHz]

Figure 6.8: Relative residual $\frac{r_n}{r_1}$ as a function of the iteration number for T0BC, OO2BC and OO0D transmission conditions using Krylov solver.

Discussion It can be seen in Figure 6.8a and 6.8b that at both frequencies, OO0D converges faster than T0BC. This result is expected from literature [6]. However in both cases, OO2BC has the best convergence. These results are consistent with the prediction on the convergence rates of the two-subdomain model problem treated in Section 2.4 and in particular with Figure 2.3a and 2.3b. Indeed, for bad conductors (*i.e.* $\frac{\tilde{\sigma}}{\omega \mathrm{Re}(\tilde{\epsilon})} \ll 1$), the model problem clued that OO2BC performs better while no advantage for OO0D or OO2BC can be concluded for dielectrics (*i.e.* $\frac{\tilde{\sigma}}{\omega \mathrm{Re}(\tilde{\epsilon})} \approx 1$). Increasing the ratio $\frac{\tilde{\sigma}}{\omega \mathrm{Re}(\tilde{\epsilon})}$ would probably give the advantage to OO0D.
Due to the frequency range and the tissues involved in this work, OO2BC is the transmissions condition that is kept.

As a conclusion, Krylov solver is always better than the Jacobi solver. In literature, Jacobi scheme is actually only used to derive convergence rates as in Section 2.4. Moreover, for the application that is considered in Chapter 8, OO2BC transmission conditions can be used indifferently in empty space and in conductive media.

Chapter 7

Numerical validation of optimization

In the preceding chapter, parameters to solve Eqs.(4.5) using domain decomposition methods and finite elements are discussed. In this chapter, this solution technique is applied sequentially for each elementary source of an array in order to compute the matrices introduced in Chapter 5. Once the matrices are known, some of their properties are checked and different optimization problems are solved.

Model Problem IV To be able to validate the results, a model problem is once again defined. This model problem consists of a 2-source array in nearly empty space. Indeed two conductive cylinders with the electrical properties of muscles (*cfr* Chapter 8) are added above and below the array. Parameters are different from Model Problem III as physical dimensions are now considered. The geometry of this model problem is given in Figure 6.1d while the numerical parameters are summarized in Table C.1 (*cfr* Appendix). The goal of this model problem is to maximize the power radiated through a surface S_r defined as

$$S_r\left(\theta_r, \Delta\theta_r\right) \triangleq \left\{ \{\rho, \theta\} \in \{\rho_r\} \times \left[\theta_r - \frac{\Delta\theta_r}{2}, \theta_r + \frac{\Delta\theta_r}{2} \right] \right\} \tag{7.1}$$

while keeping the injected power P_s constant and equal to P_0 (*cfr* Chapter 5, Eqs(5.17)).
To do so, first both elementary solutions u_1 and u_2, corresponding to the two elementary sources of the array, are computed as the solutions of Eqs.(4.5). Then the power related matrices $\underline{\underline{P_s}}$, $\underline{\underline{R}}(S_r)$ and $\underline{\underline{L}}(V)$ are built from Eq.(5.6), Eq.(5.11) and Eq.(5.9). Generalized eigenvectors and eigenvalues related to some of these matrices are finally computed to find the extrema source weights $\underline{j}$ (*cfr* Section 5.3.1).

7.1 Total outgoing power maximization

If $\Delta\theta_r = 360°$ is considered, then the problem Eqs(5.17) reduces to maximizing the total outgoing power through a closed circle around the sources. Because of power conservation, this problem should thus be equivalent to minimizing the losses L inside the circle (because $S \triangleq \partial V = S_r$). In Section 5.3, it is shown that extrema of the radiated power R or the losses L are proportional to the generalized eigenvalues λ_n^R or λ_m^L, solutions of

$$\underline{\underline{R}}\,\underline{j_n^R} = \lambda_n^R \underline{\underline{P_s}}\,\underline{j_n^R} \quad \text{or} \quad \underline{\underline{L}}\,\underline{j_m^L} = \lambda_m^L \underline{\underline{P_s}}\,\underline{j_m^L}. \tag{7.2}$$

For the closed surface considered in this section, power conservation yields

$$1 = \lambda_n^R + \lambda_m^L \quad \text{and} \quad \underline{j_n^R} = \underline{j_m^L}. \tag{7.3}$$

The short rigorous proof of this statement is given in Appendix B.8. Intuitively, it simply means that minimizing the losses inside the circle is equivalent to maximizing the power radiated outside the circle, thus the losses-minimizing weights vector equals the radiation-maximizing weights. In the following paragraphs, properties of the power matrices are verified numerically then the result of the optimization is validated from intuition.

Power matrices To check that the power related matrices $\underline{\underline{P_s}}$, $\underline{\underline{R}}$ and $\underline{\underline{L}}$ are computed correctly, the complex submatrices $\underline{\underline{P_s^b}}$, $\underline{\underline{R^b}}$ and $\underline{\underline{L^b}}$ are analysed. In the simple model example considered in this section, these matrices are given by

$$\underline{\underline{P_s^b}} = \begin{pmatrix} -8842.8 - i1622.3 & -6117.85 - i5241.92 \\ -6117.85 - i5241.92 & -8842.86 - i1622.07 \end{pmatrix}, \tag{7.4}$$

$$\underline{\underline{R^b}} = \begin{pmatrix} -380.143 + i4423.3 & -160.337 + i1866.27 \\ -160.429 + i1866.68 & -379.605 + i4423.43 \end{pmatrix}, \tag{7.5}$$

$$\underline{\underline{L^b}} = \begin{pmatrix} 19022.6 - i4390.61 & 18366.4 - i4238.99 \\ 18366.3 - i4239.31 & 19022.3 - i4390.54 \end{pmatrix}. \tag{7.6}$$

As expected because of the symmetry of Model Problem IV, the diagonal entries are nearly equal and the matrices are nearly symmetric.

Power conservation To check that the power is conserved despite numerical errors[1], one test is to check that $\underline{\underline{P_s}} - \underline{\underline{L}} - \underline{\underline{R}}$ is skew symmetric. For the model problem considered here, the relative symmetric part of this matrix is

$$\frac{\left(\underline{\underline{P_s}} - \underline{\underline{R}} - \underline{\underline{L}}\right)^T + \left(\underline{\underline{P_s}} - \underline{\underline{R}} - \underline{\underline{L}}\right)}{2\min\left(\underline{\underline{P_s}}, \underline{\underline{R}}, \underline{\underline{L}}\right)} = \begin{pmatrix} 1.31594 & 1.30961 & 0.0 & 0.0103775 \\ 1.30932 & 1.31553 & -0.0103715 & 0.0 \\ 0.0 & -0.0103775 & 1.31594 & 1.30961 \\ 0.0103715 & 0.0 & 1.30932 & 1.31553 \end{pmatrix} [\%] \tag{7.7}$$

and can thus be considered as skew symmetric. Then to have an insight of how precise are the eigenvectors and eigenvalues, Eq.(7.3) can be verified:

$$\frac{1 - \lambda^R - \lambda^L}{\min\left(\lambda^R, \lambda^L\right)} = 0.65,\ 0.66,\ 11\,[\%] \tag{7.8}$$

$$\frac{\underline{j^R} - \underline{j^L}}{\min\left(\left\|\underline{j^R}\right\|, \left\|\underline{j^L}\right\|\right)} = \begin{pmatrix} 3 \times 10^{-4} \\ -3 \times 10^{-4} \\ -8 \times 10^{-4} \end{pmatrix}, \begin{pmatrix} 7 \times 10^{-4} \\ 10^{-18} \\ -2.5 \times 10^{-8} \end{pmatrix}, \begin{pmatrix} 5 \times 10^{-5} \\ 5 \times 10^{-5} \\ 2 \times 10^{-4} \end{pmatrix} [\%]. \tag{7.9}$$

[1] One should be careful that a bad boundary condition will not produce any error as far as power conservation is concerned. Indeed while this approximation tries to mimic the unbounded solution, it is nevertheless compatible with power conservation.

The high relative error on the last eigenvalue is a consequence of the small value of the losses. Computing the absolute error shows that it is nearly the same for the three generalized eigenvalue pairs.

Radiation patterns The largest radiation eigenvalue λ_M^R, the corresponding smallest losses eigenvalue λ_m^L and their common eigenvector $\underline{j_M^R} = \underline{j_m^L}$ are[2]

$$\lambda_M^R = 93.8\,[\%], \quad \lambda_m^L = 5.6\,[\%] \quad \text{and} \quad \underline{j_M^R} A_\delta = \begin{pmatrix} -19.1 \\ 19.1 \\ 0 \end{pmatrix} \text{ [mA]}. \tag{7.10}$$

From Section 4.2, the far field radiation pattern associated with these weights is known to minimize the radiated field above and under the arrays. This results is actually expected as losses occur in the cylinders placed above and below the array.
A comparison of the incident far field (*i.e.* the far field computed with Eq.(4.20), which corresponds to the far field obtained with a 2-point source array in empty space) and the total (pseudo-)[3]far field (obtained through the numerical simulation) is given in Figure 7.1a.

Similarly, the smallest radiation eigenvalue λ_m^R, the corresponding largest losses eigenvalue λ_M^L and their common eigenvector $\underline{j_m^R} = \underline{j_M^L}$ are

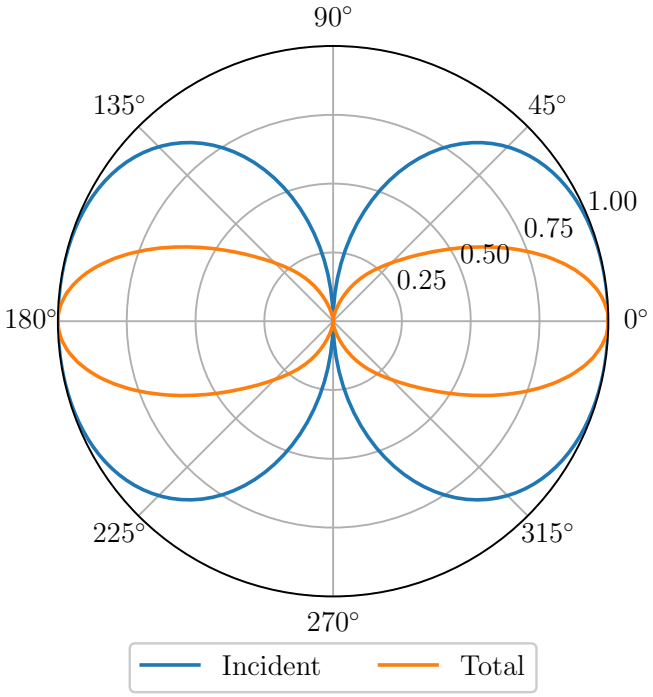
$$\lambda_m^R = 42\,[\%], \quad \lambda_M^L = 57.7\,[\%] \quad \text{and} \quad \underline{j_m^R} A_\delta = \begin{pmatrix} 8.2 \\ 8.2 \\ 0 \end{pmatrix} \text{ [mA]}. \tag{7.11}$$

From section 4.2, the far field radiation pattern associated with these weights is also known to maximize the radiated field above and under the array. This result is once again expected as losses occur in the cylinders.
A comparison of the incident far field and the total (pseudo-)far field is given in Figure 7.1b.

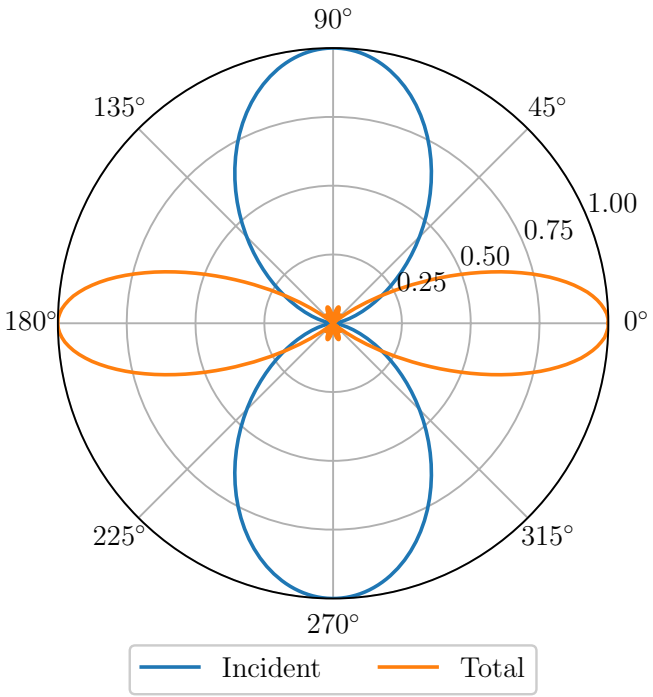
[2]While the eigenvectors have units [A/m^2] and are current densities, the quantity that is physically relevant is either the *real* current density $j_a = j_p w_p$ or the total current on the source

$$\int j_a \, dV = \int j_p w_p dV = j_p A_\delta.$$

[3]This pseudo-far field is simply obtained by computing the absolute value of the solution on the radiation boundary. This technique is correct if the radiation boundary is electrically far (*i.e.* $k_0 z \gg 1$) which is not really the case. A more precise far field could be obtained using near field to far field transformations.



(a) Radiation maximizing weights $\underline{j}_M^R$.



(b) Radiation minimizing weights $\underline{j}_m^R$.

Figure 7.1: Comparison of the incident far field and the total (pseudo-)far field for a 2-source array with optimal weights. S_r $(\theta_r = 0°, \Delta\theta_r = 180°)$.

7.2 Directive outgoing power maximization

If $\Delta\theta_r < 360^\circ$ is considered, then the radiation surface is not closed anymore and the power conservation properties Eqs(7.3) are therefore not valid anymore because $S \triangleq \partial V \neq S_r$. In this section two such problems are considered for two different radiation surfaces S_r.

$\theta_r = 0^\circ$ and $\Delta\theta_r = 180^\circ$

This case leads to

$$\lambda_M^R = 64.7\,[\%] \quad \text{and} \quad \underline{j_M^R} A_\delta = \begin{pmatrix} 12 \\ -20.1 \\ -3.7 \end{pmatrix} \text{ [mA]}, \tag{7.12}$$

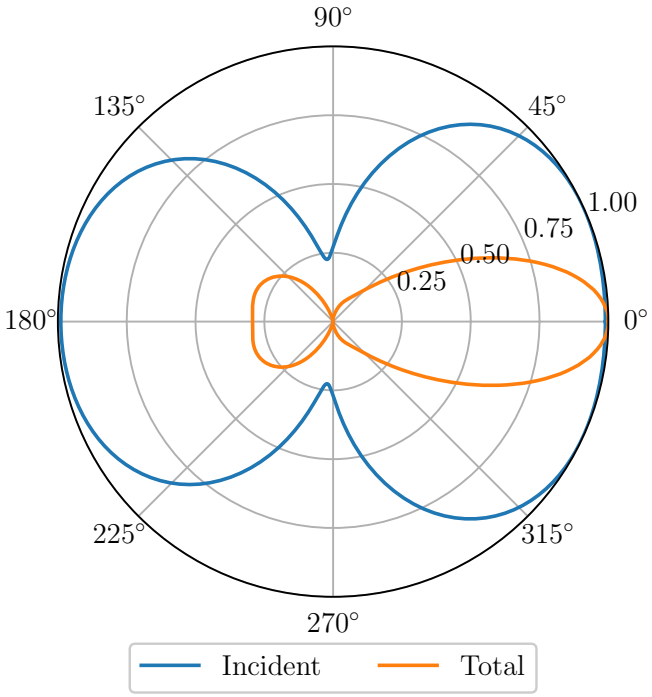
and the corresponding radiation patterns are given in Figure 7.2a.
It is interesting to highlight that the total field is very different from the preceding case while the incident field is similar (*cfr* Figure 7.1a). It is hard to guess intuitively the pair of incident and total fields of Figure 7.2a because it is hard to take into account the destructive interferences that appear on the left side of the array because of the two cylinders.

$\theta_r = 75^\circ$ and $\Delta\theta_r = 60^\circ$

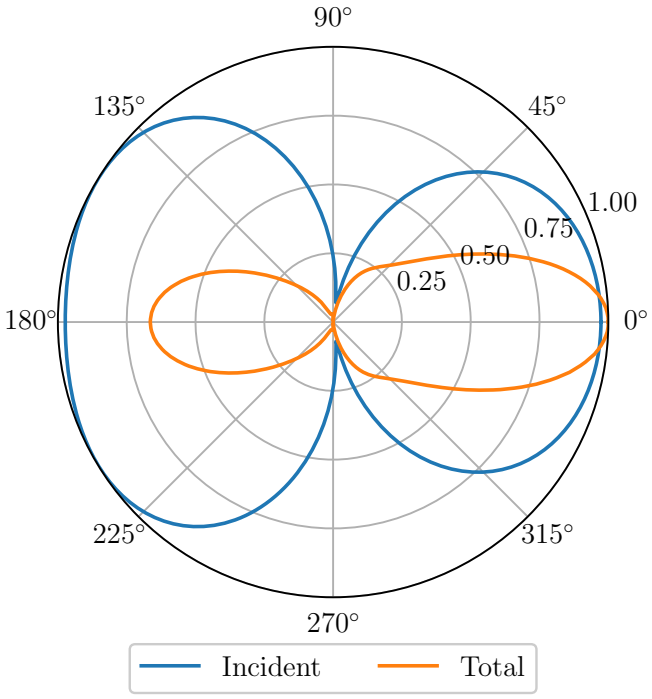
For these parameters, one finds then one finds

$$\lambda_M^R = 1.3\,[\%] \quad \text{and} \quad \underline{j_M^R} A_\delta = \begin{pmatrix} 14.1 \\ -18.2 \\ 7.1 \end{pmatrix} \text{ [mA]}. \tag{7.13}$$

The corresponding radiation pattern is given in Figure 7.2b.
As can be seen, the 2-source array have actually few impact on its radiation pattern because of the cylinders. These cylinders are very good conductors and thus allow nearly no radiation through or near them such that in this configuration it is impossible to have an array whose principal radiation direction is at $\theta = 90^\circ$. While one can maximize the radiation at $\theta = 90^\circ$ or $\theta = 75^\circ$ in this case, it is impossible to tilt the radiation pattern in presence of good conductors because they behave like shields in some directions.
This result is actually well illustrated in Figure 7.1a, 7.1b, 7.2a and 7.2b as the main radiation directions are always $\theta = 0^\circ$ or $\theta = 180^\circ$.



(a) $\theta_r = 0°, \Delta\theta_r = 180°$.



(b) $\theta_r = 75°, \Delta\theta_r = 60°$.

Figure 7.2: Comparison of the incident far field and the total (pseudo-)far field for a 2-source array with radiation maximizing weights $\underline{j}_M^R$.

Chapter 8

Power transfer through biological tissues

Now that the operation flow has been presented in the previous chapters, application to a more realistic problem is considered. The problem that is tackled in this chapter is related to power or information transfer through biological tissues and in particular, to wireless intra-body sensors. Such sensors are very useful in medicine as they provide patient mobility and continuous mesurement of critical quantities. However their power usage must be monitored carefully as their size and position make them hard to charge. Moreover excessive electromagnetic field exposure can hurt biological tissue. The aim of this chapter is thus to apply a new tool based on domain decomposition methods for optimizing the shape of antennas. The novelty of this technique is that it takes into account the entire electrical and geometrical complexity of the human body, that it operates in the frequency domain and that it can be run in parallel on several computers without any modifications. Examples of sensors that could use this technique to be optimized can for example be found in [20] or [25].

8.1 Biological tissues model

As all the developments of this work are conducted for a two dimensional scalar case, so does this application. This model does not aim to reproduce exactly the reality but rather aims to be as complex as reality, as far as numerical simulation is concerned. There are two fundamental components for this model: geometrical characteristic on the one hand and electrical properties on the other hand. Both are detailed in the following two paragraphs.

Geometrical characteristics The geometry of the model was created using color axial anatomical images from the Visible Human Project sponsored by the USA National Library of Medicine [29]. The particular image that was used is given in Figure C.3 (*cfr* Appendix) while the edges of the two dimensional finite element model are given in Figure 8.1. Five biological tissues have been modeled:

- the heart
- the lunges
- the ribcage
- some muscles
- and a layer of fat.

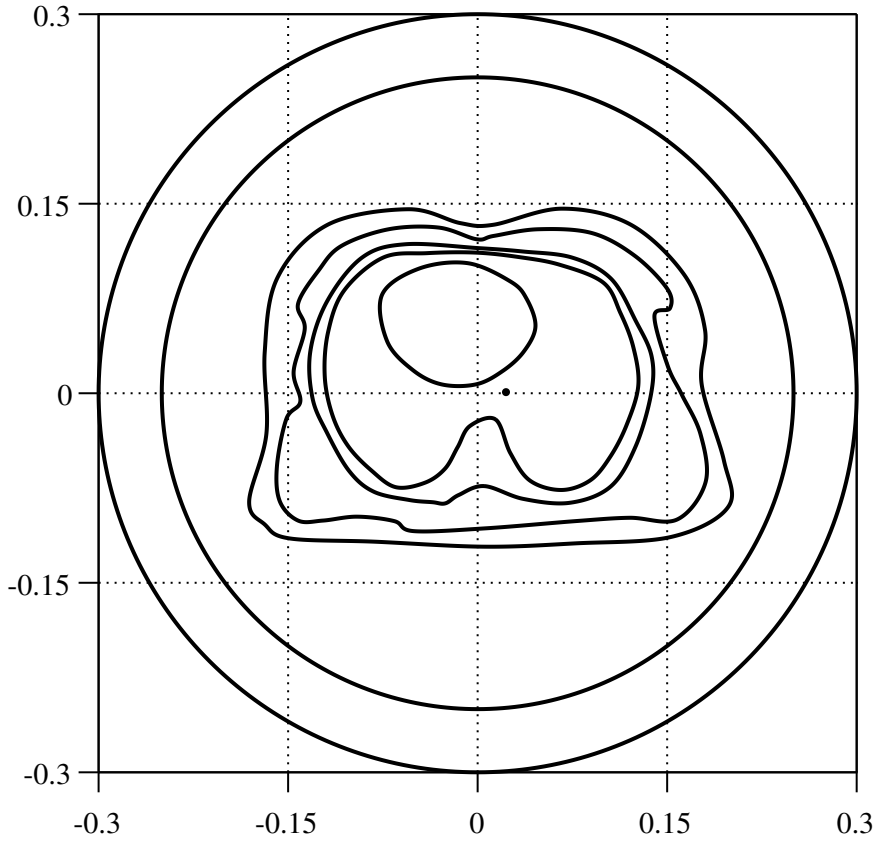


Figure 8.1: Edges of the finite element human body model considered in this work. From the center to the exterior, surfaces represent: the heart, the lunges, the ribcage, muscles, a layer of fat and empty space (twice, the inner line is used to define the radiation surface S_r). Dimensions are given in meters.

Electrical properties Each one of these media is considered non-magnetic, *i.e.* $\mu = \mu_0$ but they all have a different complex permittivity. The complex permittivity model used in this work is a summation of four Cole-Cole dispersions and a conductivity term [10], *i.e.*

$$\tilde{\epsilon}_r = \epsilon_\infty + \sum_{n=1}^{4} \frac{\Delta\epsilon_n}{1 + (i\omega\tau_n)^{(1-\alpha_n)}} + \frac{\sigma}{i\omega\epsilon_0}. \tag{8.1}$$

Each one of the caracteristic relaxation time τ_n has an influence in a limited range of frequencies. Indeed if $\omega\tau_n \gg 1$ then $\Delta\epsilon_n$ looses its influence. Therefore when working at frequencies around $f = 3$ [GHz], the relaxation time τ_3 and τ_4 have a very little influence (*cfr* Table C.2 in Appendix). Nevertheless, the four dispersions are still considered for generality. The parameters of this four Cole-Cole dispersions have been determined empirically by [10] and are reproduced in Table C.2 (*cfr* Appendix) for the five tissues that are modeled here. Using Eq.(8.1), the wavelength in the media can then be computed as

$$\lambda = c\sqrt{\mu_0\epsilon_0 \mathrm{Re}\left(\tilde{\epsilon}_r\right)} \tag{8.2}$$

and the equivalent conductivity, governing the losses, as

$$\tilde{\sigma} = -\mathrm{Im}\left(\omega\epsilon_0\tilde{\epsilon}_r\right). \tag{8.3}$$

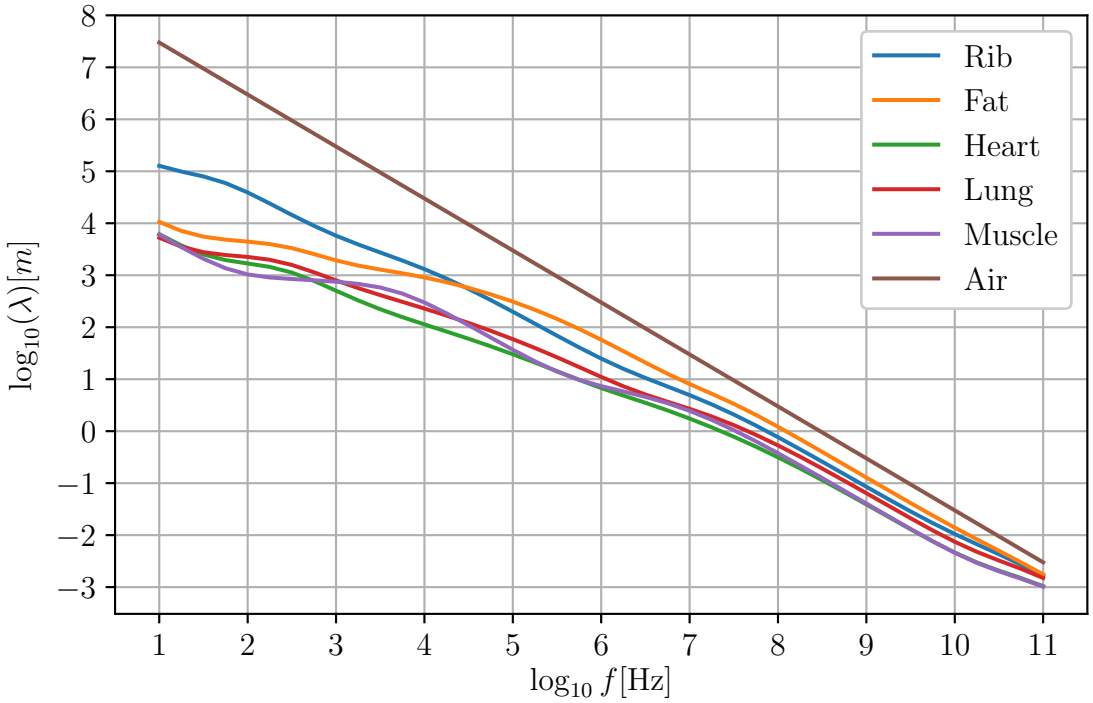


Figure 8.2: Wavelength λ for the different tissues considered in the human body model.

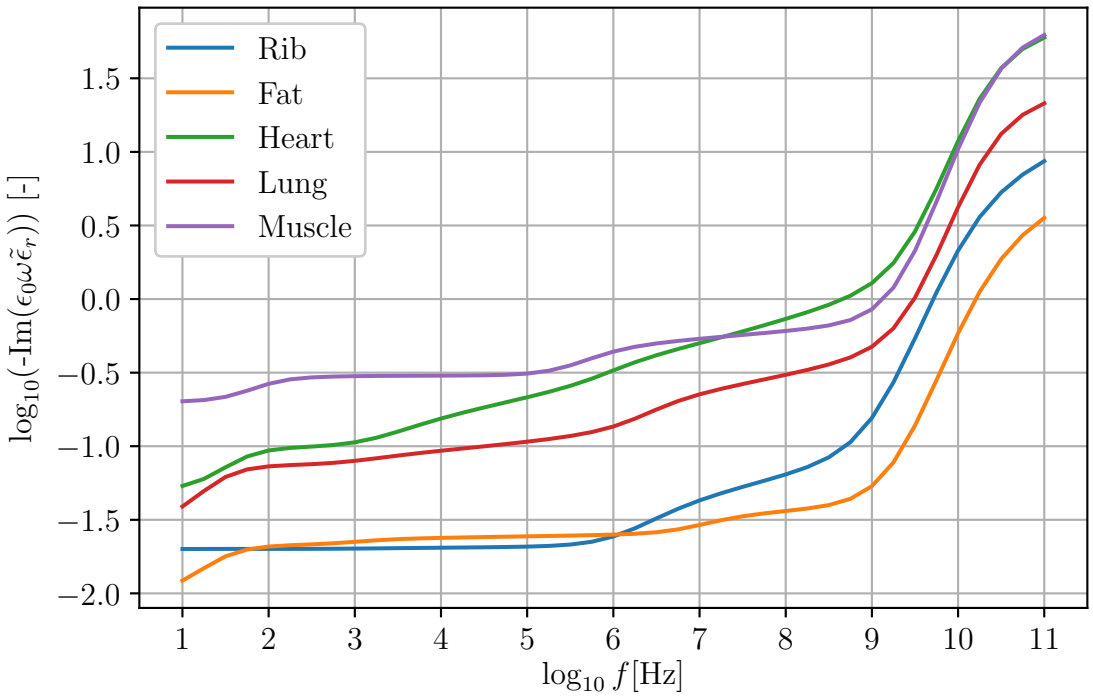


Figure 8.3: Equivalent conductivity $\tilde{\sigma}$ for the different tissues considered in the human body model.

The wavelength and the equivalent conductivity are given in Figure 8.2 and 8.3. As can be seen, in the gigahertz range, the wavelength in the muscles and the heart is up to ten times smaller than in the air. Moreover, the losses are greater in the muscles and the heart such that at any frequency, it is harder to transfer power through these media than through the others. The shielding behaviour of muscles has already been obsverved in Section 7.2. Power transfer is also very difficult at frequencies above a few GigaHertz as the losses in all media strongly increase. The real part of the relative permittivity is given in Figure C.4 (*cfr Appendix)*. The conductor ratio $\frac{\tilde{\sigma}}{\omega\epsilon_0 \mathrm{Re}(\tilde{\epsilon}_r)}$ is also given in Figure C.5 (*cfr* Appendix). This ratio allows to characterize media (*cfr* Section 1.1) and to chose the transmission condition accordingly, as explained in Section 2.4.

8.2 Specific absorption rate based constraints

Critical quantities in the study of biological effects of electromagnetic fields are the *specific absorption rate* [W/kg] defined as

$$\text{SAR} \triangleq \frac{\tilde{\sigma} \left\| \boldsymbol{e} \right\|^2}{2\rho} \tag{8.4}$$

and its *whole body average* defined as

$$\text{WBA-SAR} \triangleq \frac{1}{V} \int_V \frac{\tilde{\sigma} \left\| \boldsymbol{e} \right\|^2}{2\rho} \, d\boldsymbol{x} \tag{8.5}$$

where ρ [kg/m^3] is the *density* of the tissue and V the volume of the body.
From a biological point of view, the harmful mechanism is actually the temperature rise due to dissipated power. Under the assumption that heat conduction is negligible, the specific absorption rate is linked to the local temperature increase by

$$T = T_0 + \frac{1}{c} \text{SAR} \times (t - t_0) \tag{8.6}$$

where c [J/K/kg] is the *specific heat capacity*. Proof of this link and validity of the assumption are given in Appendix B.9. In the International Commission on Non-Ionizing Radiation Protection (ICNIRP) guidelines [17] and the IEEE standard [16], the WBA-SAR has a basic restriction of $\text{WBA-MAX}_{\text{max}} = 0.08$ [W/kg] for general public exposure. Using a specific heat typical of muscle [22], *i.e.* $c = 0.9$ [cal/g/K] $= 3766$ [J/kg/K] and the maximal WBA-SAR, the rate of temperature increase is then

$$\frac{\text{SAR}}{c} = 2 \times 10^{-5} \text{ [K/s]} = 0.08 \text{ [K/h]}. \tag{8.7}$$

The finite element framework used here allows to compute the exact temperature field by solving the bio-heat equation[1] [14]. The SAR could also be computed for each tissue simply by dividing the loss term by the corresponding density ρ. However, it has been chosen for simplicity to consider a restriction similar to those on the WBA-SAR because they can be linked to the losses L inside the

[1]The bio-heat equation is another formulation of the heat equation Eq.(B.83) specified for heat transfer inside biological tissues

body, *i.e.*

$$\begin{aligned}\text{WBA-SAR}_{\max} \geq \text{WBA-SAR} &= \frac{1}{V}\int_V \frac{\tilde{\sigma}\left\|\boldsymbol{e}\right\|^2}{2\rho}\,d\boldsymbol{x} \\ &\approx \frac{1}{V\rho_m}\frac{1}{2}\int_V \tilde{\sigma}\left\|\boldsymbol{e}\right\|^2\,d\boldsymbol{x} \qquad (8.8)\\ &= \frac{L}{V\rho_m}. \qquad (8.9)\end{aligned}$$

In the optimization problem that is solved for this application, the constraint is thus

$$L = L_0 = V\rho_m \text{WBA-SAR}_{\max}. \qquad (8.10)$$

For biological tissue, the density is typically 1000 [kg/m^3] [22] and from Figure 8.1 the volume of the body can be estimated as $V = 0.3 \times 0.3 = 9 \times 10^{-2}$ [m^2] such that the constraint is taken as

$$L = L_0 = V\rho_m \text{WBA-SAR}_{\max} = 7.2 \text{ [W/m]}. \qquad (8.11)$$

8.3 Phased array configuration

The considered phased arrays are centered around the black dot located in the lunges in Figure 8.1. They consist of either $N_p = 1, 2, 5$ or 9 sources following the pattern of Figure 8.4. The sources spacing d is taken as $d = \lambda_m/2$ where λ_m is the reference wavelength inside the tissues. This reference wavelength is chosen from Figure 8.2 as $\lambda_m = \lambda/10 \approx 1$ [cm], *i.e.* as the smallest wavelength. The idea behind this array is to be as symmetric as possible not to favor any direction *a priori*. Then once the optimization is performed, sources with very small weights could possibly be deleted or moved. The spacing between the sources is taken such that useful interference could appear between sources in any media. The value $d = \lambda_m/2$ is chosen from the reasoning of Section 4.2. Ideally, the maximal spacing between two sources should thus be half the maximal wavelength.

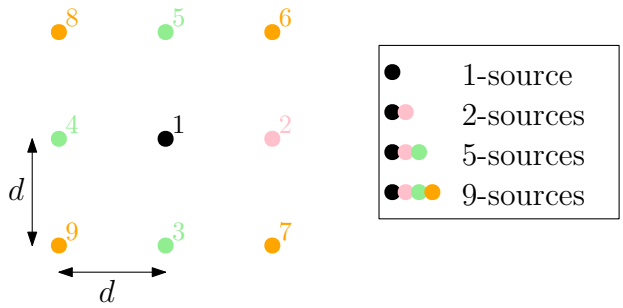


Figure 8.4: Sources relative placement around their center. The center is denoted by a black dot in Figure 8.1.

8.4 Optimization

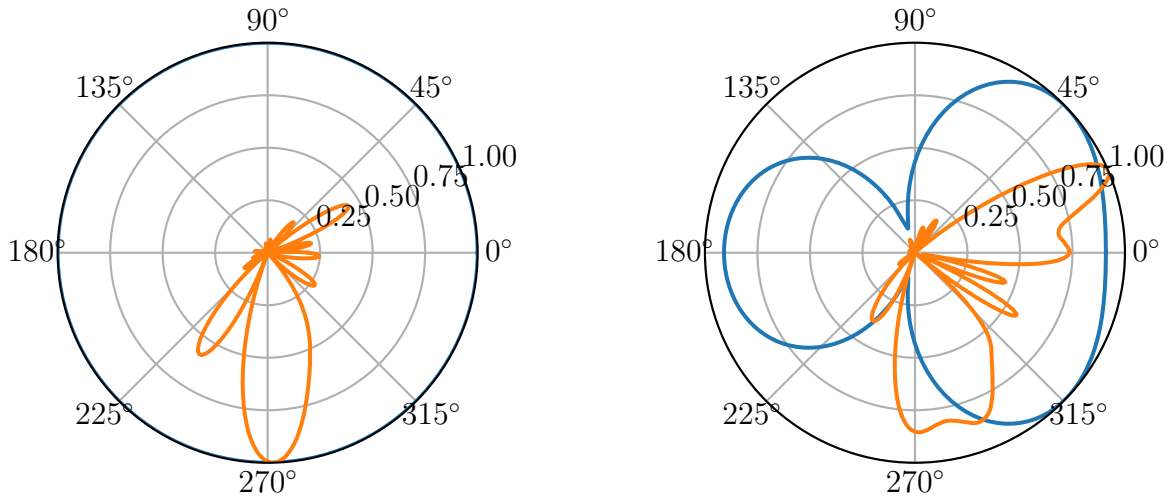
Following the procedure of Chapter 7, the cost function for this application is chosen to be the radiated power R through S_r with $\theta_r = 45°$ and $\Delta\theta_r = 45°$. The frequency used is once again $f = 3$ [GHz]. The parameters for the numerical simulation are summarized in Table C.1 (*cfr* Appendix). It is important to highlight that only $n_0 = n_m = 12$ points per wavelengths (in empty space and in the media) are used for a total of approximately 5×10^5 elements. A much finer mesh could easily be used if the simulations were made parallel on a computer cluster.

Results The relative radiated power through the surface S_r for $N_p = 1, 2, 5$ and 9 sources is given in Table 8.1. The corresponding source weights are given in Table C.3 (*cfr* Appendix) and the (pseudo-)far field radiation patterns are given in Figure 8.5. The real part of the total field u is also given in Figure 8.6 for $N_p = 1$ and $N_p = 9$.

	$N_p = 1$	$N_p = 2$	$N_p = 5$	$N_p = 9$
λ [%]	2.24×10^{-6}	6.261×10^{-6}	8.83×10^{-6}	16.1×10^{-6}

Table 8.1: Relative radiated power through the surface $S_r(\theta_r = 45°, \Delta\theta_r = 45°)$ for $N_p = 1, 2, 5$ and 9 sources. The total power is obtained by multiplying λ by L_0.

Discussion As can be seen from Figure 8.5a and 8.6a, for the 1-source case the muscle and the heart act like very efficient shields and it is thus harder to radiate in some directions. The radiation in directions between $\theta = 45°$ and $\theta = 225°$ is extremely weak due to the relative position of the sources and the heart. Along directions between $\theta = 225°$ and $\theta = 45°$, the radiation is efficient only when the muscle layer is thin. In particular, it seems that there is a window for radiation at $\theta = 270°$. These results on the shielding properties of heart and muscles are expected from the results on Model Problem IV.
While it is hard to tune the direction of the main lobe, this does not mean that the performance in the direction of interest is not increased. Indeed, the relative radiated power (that is tabulated in Table 8.1) increases from $N_p = 1$ to $N_p = 9$. The output power in the desire direction is actually multiplied by 8, while the loss and the injected power are kept constant because $P_s = L_0 + R$ with $R = \lambda L_0 \ll L_0$.
Furthermore, Figure 8.5 and 8.6 also show that increasing the number of sources from $N_p = 1$ to $N_p = 9$ allows to tilt to main lobe from $\theta = 270°$ to $\theta = 0°$ while the window at $\theta = 270°$ seems initially much more open (based on the isotropic radiation pattern for $N_p = 1$).
One should be careful that in Figure 8.5, the curves are normalized to set their maximum values to 1. It is therefore impossible to compare their amplitudes for different number of sources N_p.



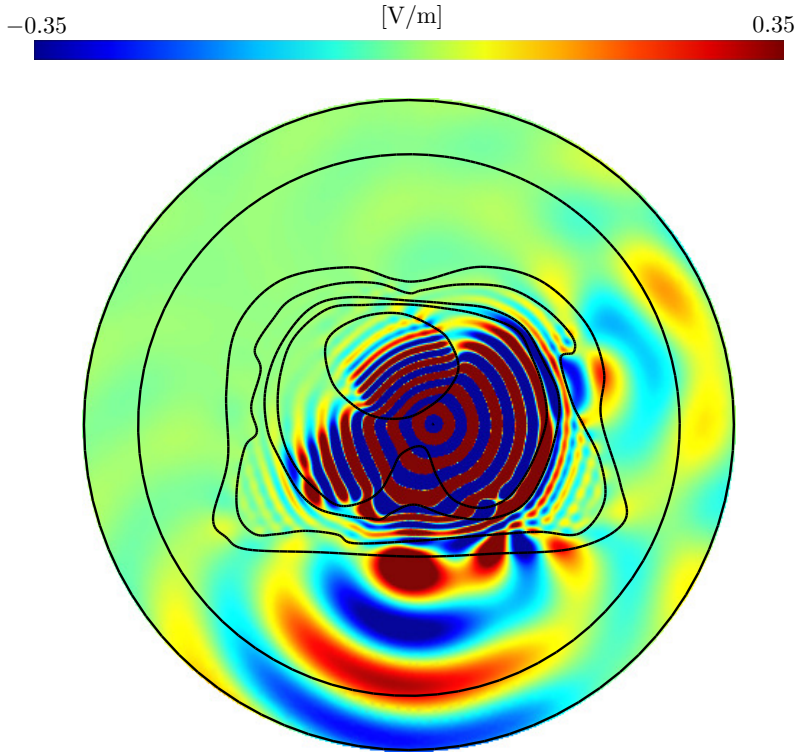
(a) $N_p = 1$

(b) $N_p = 2$

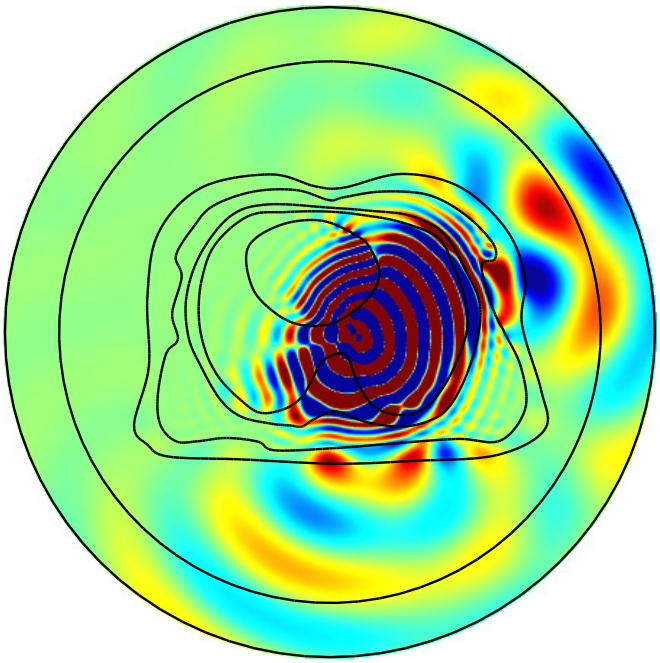
(c) $N_p = 5$

(d) $N_p = 9$

Figure 8.5: Incident (blue) and total (orange) (pseudo-)far field radiation pattern for the optimal sources weights of Table C.3, for different number of sources N_p. The maximum value of each curved is set to one.



(a) $N_p = 1$



(a) $N_p = 9$

Figure 8.6: Real part of the total field $\mathrm{Re}(u)$ [V/m] for the optimal sources weights of Table C.3, for $N_p = 1$ and 9 sources.

Conclusion

In this work, a particular kind of domain decomposition methods, the non-overlapping Optimized Schwarz algorithms, is developed for the frequency domain Maxwell's equations for transverse fields. This solution technique is then used to optimize the radiation properties of phased arrays in non-homogeneous media and in particular in biological tissues.

First of all, a formulation of these domain decomposition methods as solvers is explained. Then different variants, *i.e.* different transmission conditions, for common media in electromagnetism are discussed and compared on a model problem. Finally, the finite element discretization of the algorithms is explained in the particular case of Lagrange $\mathbb{P}_1$ elements.
Phased arrays are then defined and optimization problems based on power budget are introduced. A solution technique which requires the use of non-overlapping Schwarz algorithms is derived.
Each step of the solution technique is then validated on model problems for which analytical solution is known. Error are therefore also estimated. The model problems are chosen to be close to the final realistic application: the optimization of the radiated power in a particular direction for a phased array surrounded by biological tissues. The optimization allows to multiply the radiated power by 8, proving the efficiency of the method.

The realistic application considered at the end of this work is promising and let believe that antenna designer could benefit from this new method, especially because it speeds up easily when using computer clusters. It would be interesting to generalize the results developed here to three dimensional non transverse fields to fully exploit the possibilities of the method. Once this generalization is done, antennas could be optimized for full-complexity media.

Bibliography

[1] G. Allaire. *Analyse numérique et optimisation*. Editions de l'Ecole polytechnique, 2005.

[2] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A Fresh Approach to Numerical Computing. *Soc. Ind. Appl. Math.*, 59(1), 2017.

[3] M. E. Bouajaji, X. Antoine, and C. Geuzaine. Approximate local magnetic-to-electric surface operators for time-harmonic Maxwell ' s equations. *J. Comput. Phys.*, 279 (15):pp.241–260, 2014.

[4] Y. Boubendir, X. Antoine, and C. Geuzaine. A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation. *J. Comput. Phys.*, 231(2):262–280, 2012.

[5] R. T. Couto. Green's functions for the wave, Helmholtz and Poisson equations in a two-dimensional boundless domain. *Rev. Bras. Ensino Fis.*, 1304(35):173–176, 2013.

[6] V. Dolean, M. E. Bouajaji, M. J. Gander, and S. Lanteri. Optimized Schwarz methods for Maxwell's equations with non-zero electric conductivity. *Lect. Notes Comput. Sci. Eng.*, 78 LNCSE(1):269–276, 2010.

[7] V. Dolean, P. Jolivet, and F. Nataf. *An introduction to domain decomposition methods*. Society for Industrial and Applied Mathematics, 2015.

[8] P. Dular, C. Geuzaine, F. Henrotte, and W. Legros. A general environment for the treatment of discrete problems and its application to the finite element method. *IEEE Trans. Magn.*, 34(5):3395–3398, 1998.

[9] O. G. Ernst and M. J. Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. *Lect. Notes Comput. Sci. Eng.*, 83:325–363, 2012.

[10] S. Gabriel, R. W. Lau, and C. Gabriel. The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues. *Phys. Med. Biol.*, 41(11):2271–2293, 1996.

[11] M. J. Gander, F. Magoulès, and F. Nataf. Optimized Schwarz method without overlap for the Helmholtz equation. *SIAM J. Sci. Comput.*, 24(1):38–60, jan 2001.

[12] J. Gardner and R. E. Collin. An Accurate Closed-Form Approximate Representation for the Hankel Function of the Second Kind. *IEEE Trans. Antennas Propag.*, 48(11):1699–2, 2000.

[13] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre-and post-processing facilities. *Int. J. Numer. Methods Eng. 79(11)*, 0:1309–1331, 2009.

[14] A. Hirata, I. Laakso, T. Oizumi, R. Hanatani, K. H. Chan, and J. Wiart. The relationship between specific absorption rate and temperature elevation in anatomically based human body models for plane wave exposure from 30 MHz to 6 GHz. *Phys. Med. Biol.*, 58(4):903–921, 2013.

[15] B. Hosseini, N. Nigam, and J. M. Stockie. On regularizations of the Dirac delta distribution. *J. Comput. Phys.*, 305:423–447, 2016.

[16] IEEE. Standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3kHz to 300 GHz. *IEEE Std C95.1-2005 (Revision IEEE Std C95.1-1991)*, pages 1–238, 2006.

[17] International Commission on Non-Ionizing Radiation Protection. Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz). *Health Phys.*, 97(3):257–258, sep 2009.

[18] I. C. Ipsen and C. D. Meyer. The idea behind Krylov methods. *Am. Math. Mon.*, 105(10):889–899, 1998.

[19] J.-M. Jin. *The Finite Element method in Electromagnetics*. Wiley-IEEE Press, 2014.

[20] T. Karacolak, A. Z. Hood, and E. Topsakal. Design of a dual-band implantable antenna and development of skin mimicking gels for continuous glucose monitoring. *IEEE Trans. Microw. Theory Tech.*, 56(4):1001–1008, 2008.

[21] J. D. Kraus and R. J. Marhefka. *Antennas for all applications*. McGraw-Hill, 2001.

[22] M. Lipkin and J. D. Hardy. Measurement of some thermal properties of human tissues. *J. Appl. Physiol.*, 7(2):212–7, 1954.

[23] F. Nataf. Absorbing boundary conditions and perfectly matched layers in wave propagation problems. In *Direct and Inverse problems in Wave Propagation and Applications*, volume 14 of *Radon Ser. Comput. Appl. Math.*, pages 219–231. de Gruyter, 2013.

[24] Y. Saad and M. H. Schultz. GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM J. Sci. Stat. Comput.*, 7(3):856–869, 1986.

[25] M. L. Scarpello, D. Kurup, H. Rogier, D. V. Ginste, F. Axisa, J. Vanfleteren, W. Joseph, L. Martens, and G. Vermeeren. Design of an implantable slot dipole conformal flexible antenna for biomedical applications. *IEEE Trans. Antennas Propag.*, 59(10):3556–3564, 2011.

[26] S. Silver. *Microwave antenna theory and design*. McGraw-Hill, 1949.

[27] C. T. Tai. Complementary reciprocity theorems in electromagnetic theory. *IEEE Trans. Antennas Propag.*, 40(6):675–681, 1992.

[28] B. Thierry, A. Vion, S. Tournier, M. El Bouajaji, D. Colignon, N. Marsic, X. Antoine, and C. Geuzaine. GetDDM: An open framework for testing optimized Schwarz methods for time-harmonic wave problems. *Comput. Phys. Commun.*, 203:309–330, 2016.

[29] U.S. National Library of Medicine. Visible Human Project. `https://www.nlm.nih.gov/research/visible/visible_human.html`. [Online; accessed 2018-08-21].

[30] C. H. Wilcox. An expansion theorem for electromagnetic fields. *Commun. Pure Appl. Math.*, 9(2):115–134, 1956.

[31] S. Zahedi and A. K. Tornberg. Delta function approximations in level set methods by distance function extension. *J. Comput. Phys.*, 229(6):2199–2219, 2010.

Appendix A

Conventions and notations

Convention	Definition
$\triangleq$	a definition
Eq.$(x.y)$	equation from Chapter x numbered y
Eqs$((x.y))$	system of equations from Chapter x numbered y
text	first introduction of a concept/quantity called text
a^*	the conjugate of a
$\boldsymbol{a}$	a frequency domain vector
$\mathbf{a}$	a time domain vector
a	a frequency domain scalar
a	a time domain scalar
$\hat{a}$	the spatial Fourier transform of a
$\hat{\boldsymbol{a}}$	a unit vector
∂a	the boundary of set a
a^n	a quantity at the nth iteration
$\overline{a}$	the closure of a $(a \bigcup \partial a)$
$\tilde{a}$	a test function
$\underline{a}$	a vector
$\underline{\underline{a}}$	a matrix

Symbol	Name
$i = \sqrt{-1}$	imaginary number
ω	pulsation
$\mathcal{F}^{(-1)}$	(inverse) Fourier transform
$\boldsymbol{h}$	magnetic field
$\boldsymbol{e}$	electric field
$\boldsymbol{b}$	magnetic flux density
$\boldsymbol{d}$	electric displacement field
$\boldsymbol{j}$	current density
ρ	charge density
$\boldsymbol{j}_c$	conduction current density
$\boldsymbol{j}_a$	applied current density

σ	conductivity
μ	magnetic permeability
ϵ	electric permittivity
$\tilde{\epsilon}$	complex electric permittivity
$\tilde{\sigma}$	equivalent conductivity
GC	good conductor
BC	bad conductor
D	dielectric
k, k_0	wave number
$\boldsymbol{k} = (k_x, k_y)$	wave vector
u, u_{inc}, u_{scat}	TE or TM total, incident, scattered fields
f	TE or TM source term
ρ	cylindrical coordinate radius
PEC	perfect electric conductor
PMC	perfect magnetic conductor
Λ	Dirichlet to Neuman (DtN)
θ	incidence angle
κ	curvature
T0	zeroth order Taylor approximation for plane surfaces
CT0	zeroth order Taylor approximation for curved surfaces
T2	second order Taylor approximation for plane surfaces
CT2	second order Taylor approximation for cuvred surfaces
u_∞	RHS of absorbing boundary conditions
u_d	RHS of Dirichlet boundary conditions
u_n	RHS of Neumann boundary conditions
$\partial\Omega^\infty$	absorbing boundary conditions surface
$\partial\Omega^d$	Dirichlet boundary conditions surface
$\partial\Omega^n$	Neumann boundary conditions surface
$\mathcal{B}$	tangential absorbing boundary operator
g_{ij}, s_{ij}	surface unknowns/artificial sources
Γ_{ij}	common boundary between subdomains i and j
D_i	set of neighbours of subdomain i
$\mathcal{A}$, $\mathcal{A}_{ji}$, $\boldsymbol{\mathcal{A}}$	surface operator
$\mathcal{A}'$, $\mathcal{A}'_{ji}$, $\boldsymbol{\mathcal{A}'}$	linear part of surface operator
b, b_{ji}, $\boldsymbol{b}$	constant part of surface operator
$\boldsymbol{r}$	residual
ϵ_{it}	tolerance on the residual
$\mathcal{K}_n(\boldsymbol{\mathcal{A}}, \boldsymbol{b})$	Krylov subspace generated by $\boldsymbol{\mathcal{A}}$ on $\boldsymbol{b}$
ϵ_i	absolte error on subdomain i
ρ	convergence rate
$\mathcal{I}$	unitary operator
$\mathcal{S}$	tangential part of DtN based transmission condition
σ	principal symbol of $\mathcal{S}$
k_{min},k_{max}, k_+, k_-	optimized transmission condition parameters
h	characteristic mesh size
L	characteristic domain size

δ	skin depth
T0(GC-BC)	zeroth order Taylor approximation transmission condition for good-bad conductors
T2(BC-BC)	second order Taylor approximation transmission condition for good-bad conductors
OO2BC	optimized second order transmission condition for bad conductors
OO0D	optimized second order transmission condition for dielectrics
E_ν	element ν
ψ^s	surface node function
ψ^v	volume node function
N_Ω	nodes of Ω
$\underline{\underline{K}}$	stiffness matrices
$\underline{\underline{M}}$	mass matrices
N_p	source number
$\boldsymbol{w}_p$	elementary source function
j_p	source p weight
u_p	elementary solution p
$g_0(\boldsymbol{x};\boldsymbol{y})$	free space Green function
$g_0^{\text{FF}}(\boldsymbol{x};\boldsymbol{y})$	far field free space Green function
$g(\boldsymbol{x};\boldsymbol{y})$	non-homogeneous space Green function
d	source spacing
z, z_p	dimensionless radius
F	radiation pattern
A	field amplitude
$\delta(\boldsymbol{x})$	Dirac distribution
ϵ	physical point source radius
$\underline{j}$	source weights vector
P_s	injected power
R	radiated power
L	losses/dissipated power
λ	generalized eigenvalue
λ_0, λ	wavelength in empty space
λ	wavelength in complex media
λ_m	media reference wavelength
f	frequency
b	absorbing boundary radius
m	element per point source radius
CST	constant point source approximation
GSN	Gaussian point source approximation
S_r	radiation surface
θ_r	radiation surface mean azimuth
$\Delta\theta_r$	radiation surface arc length
SAR	specific absorption rate
WBA-SAR	whole body average specific absorption rate
ρ	density
T	temperature
c	specific heat capacity
k	thermal conductivity

Appendix B

Additional proofs, developments and theorems

B.1 Dirichlet to Neumann (DtN) operator for a model problem

Consider the Helmholtz equation on the semi-infinite space $\{x, y\} \in \mathbb{R}^+ \times \mathbb{R}$ for an homogeneous medium with a Sommerfeld radiation condition and a Dirichlet condition

$$\begin{cases} -\text{div } \textbf{grad } w - k^2 w = 0 & \text{in } \mathbb{R}^+ \times \mathbb{R}, \\ \lim\limits_{\rho \to \infty} \sqrt{\rho} \left(\dfrac{\partial}{\partial \rho} + jk \right) w = 0 & \\ w = f(y) & \text{on } \{x, y\} \in \{0\} \times \mathbb{R}. \end{cases} \tag{B.1}$$

Taking the Fourier transform w.r.t. y, Eqs(B.1) become

$$\begin{cases} \partial^2_{xx} \hat{w} + \left(k^2 - k_y^2\right) \hat{w} = 0 & \text{for } x > 0 \text{ and } k_y \in \mathbb{R}, \\ \lim\limits_{x \to \infty} \sqrt{x} \left(\dfrac{\partial}{\partial x} + jk \right) \hat{w} = 0, & \text{for } k_y \in \mathbb{R}, \\ \hat{w} = \hat{f}(k_y) & \text{for } x = 0 \text{ and } k_y \in \mathbb{R}. \end{cases} \tag{B.2}$$

Defining

$$\lambda^2 \triangleq k_y^2 - k^2, \tag{B.3}$$

the solution to the first equation of Eqs(B.2) is

$$\hat{w} = A \exp -\lambda x + B \exp \lambda x \tag{B.4}$$

with[1]

$$\lambda(k_y) = \begin{cases} \sqrt{k_y^2 - k_0^2} & \text{for } |k_y| \geq |k_0| \\ j\sqrt{k_0^2 - k_y^2} & \text{for } |k_y| \leq |k_0| \,. \end{cases} \tag{B.5}$$

Then applying Sommerfeld radiation condition and Dirichlet boundary condition of Eqs(B.2), the solution becomes

$$\hat{w} = \hat{f}(k_y) \exp -\lambda x \tag{B.6}$$

[1] Eqs(B.5) stands only if k is real.

which verifies

$$\hat{\boldsymbol{n}} \cdot \mathbf{grad}\ \hat{w} = \lambda \hat{w}. \tag{B.7}$$

The DtN operator Λ can then be found by taking the inverse Fourier transform w.r.t. y on both sides of Eq.(B.7) on the surface $\{0\} \times \mathbb{R}$, *i.e.*

$$\begin{aligned} \mathcal{F}^{-1}(\hat{\boldsymbol{n}} \cdot \mathbf{grad}\ \hat{w}) &= \mathcal{F}^{-1}(\lambda \hat{w}) && \text{on } \{0\} \times \mathbb{R}, \\ \hat{\boldsymbol{n}} \cdot \mathbf{grad}\ w &= \mathcal{F}^{-1}\left(\lambda \mathcal{F}(f)\right) && \text{on } \{0\} \times \mathbb{R}, \\ &= \Lambda(f) && \text{on } \{0\} \times \mathbb{R}. \end{aligned} \tag{B.8}$$

B.2 Reflection coefficient for curved boundaries

Consider the curved boundary represented in Figure B.1, the local cylindrical coordinate system (ρ, θ) associated with any point P of this boundary and the incident field

$$u_{\text{inc}} = \exp\left(-i\boldsymbol{k} \cdot \boldsymbol{\rho}\right). \tag{B.9}$$

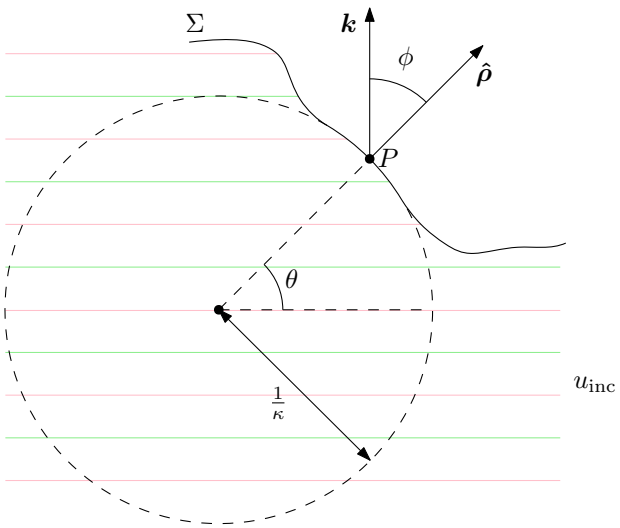


Figure B.1: Illustration of an incident plane wave on a curved boundary.

At point P, the reflected field is then

$$u_{\text{scat}} = \Gamma \exp\left(i\boldsymbol{k} \cdot \boldsymbol{\rho}\right). \tag{B.10}$$

In local cylindrical coordinates, $\hat{\boldsymbol{n}} \cdot \mathbf{grad} = \frac{\partial}{\partial \rho}$ and $\text{div}_\Sigma \mathbf{grad}_\Sigma = \kappa^2 \frac{\partial^2}{\partial^2 \theta}$. The angle between the local normal $\hat{\boldsymbol{\rho}}$ and the wave vector $\boldsymbol{k}$ is called ϕ. Because $\boldsymbol{k}$ has a constant direction, $\phi + \theta = \text{Cst}$ and therefore

$$\frac{\partial^2}{\partial^2 \theta} = \frac{\partial^2}{\partial^2 \phi}. \tag{B.11}$$

A boundary condition of the form

$$\mathcal{L}(u) = \hat{\boldsymbol{n}} \cdot \mathbf{grad}\ u + \alpha u - \beta textdiv_S igma \mathbf{grad}_\Sigma\ u = 0. \tag{B.12}$$

becomes

$$\mathcal{L}(u) = \frac{\partial u}{\partial \rho} + \alpha u - \beta \kappa^2 \frac{\partial^2}{\partial^2 \phi} = 0. \tag{B.13}$$

Inserting the total field

$$\begin{aligned} u &= \exp\left(i\boldsymbol{k}\cdot\boldsymbol{\rho}\right) + \Gamma \exp\left(-i\boldsymbol{k}\cdot\boldsymbol{\rho}\right) \\ &= \exp\left(ik_0\rho\cos\phi\right) + \Gamma \exp\left(-ik_0\rho\cos\phi\right) \end{aligned} \tag{B.14}$$

into Eq.(B.13), yields

$$\begin{aligned} &\mathcal{L}\left(u_{\text{inc}}\right) + \Gamma\mathcal{L}\left(u_{\text{scat}}\right) = 0 \\ \Rightarrow \quad &\Gamma = -\frac{\mathcal{L}\left(u_{\text{inc}}\right)}{\mathcal{L}\left(u_{\text{scat}}\right)} \\ \Rightarrow \quad &\Gamma = -\frac{u_{\text{inc}}}{u_{\text{scat}}} \frac{-ik_0\cos\phi + \alpha + \beta k_0^2 \sin^2\phi - ik_0\beta\kappa\cos\phi}{+ik_0\cos\phi + \alpha + \beta k_0^2 \sin^2\phi + ik_0\beta\kappa\cos\phi} \\ \Rightarrow \quad &\Gamma = -\frac{u_{\text{inc}}}{u_{\text{scat}}} \frac{-i\cos\phi + \frac{\alpha}{k_0} + \beta k_0 \sin^2\phi - i\beta\kappa\cos\phi}{+i\cos\phi + \frac{\alpha}{k_0} + \beta k_0 \sin^2\phi + i\beta\kappa\cos\phi}. \end{aligned} \tag{B.15}$$

The magnitude of this reflection coefficient is given in Figure B.2 for $\kappa/k_0 = 0.1$ ($\Rightarrow \lambda = 0.63R$). The bad performance of CT0 w.r.t. T0 is explained by the fact that CT0 (and CT2) are designed to absorb circular decaying waves while the performances are here evaluated for plane waves [19].

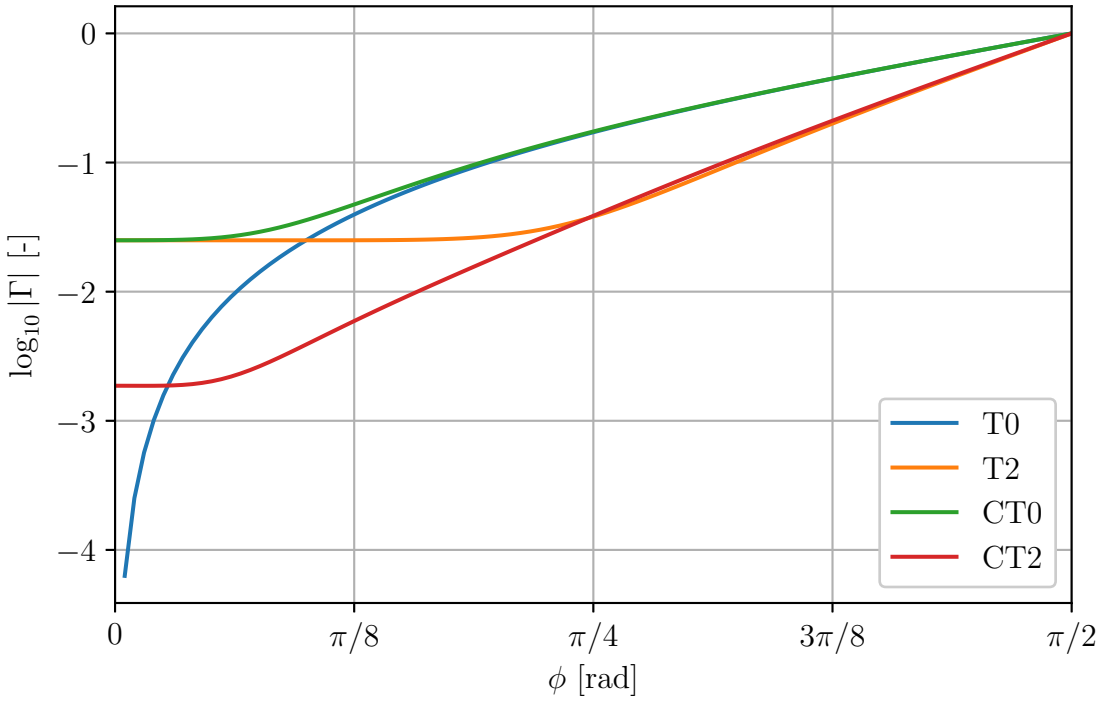


Figure B.2: Reflection coefficient Γ for T0 Eq.(1.55), T2 Eq.(1.62), CT0 Eq.(1.61) and CT2 Eq.(1.65) for curved boundaries. $\kappa/k_0 = 0.1$.

B.3 Divergence theorem

Consider Ω a compact subset of $\mathbb{R}^n$, $\partial\Omega$ its boundary and $\boldsymbol{F}$ a vector field defined on Ω. The divergence theorem then writes as

$$\int_\Omega \text{div}\ \boldsymbol{F}\, d\Omega = \int_{\partial\Omega} \boldsymbol{F}\cdot\hat{\boldsymbol{n}}\, d\partial\Omega. \tag{B.16}$$

Corollary In the particular case where $\boldsymbol{F} = g\ \textbf{grad}\ f$, Eq.(B.16) becomes

$$\int_\Omega \text{div}\,(g\ \textbf{grad}\ f)\ d\Omega = \int_{\partial\Omega} g\ \textbf{grad}\ f\cdot\hat{\boldsymbol{n}}\, d\partial\Omega$$

$$\int_\Omega (g\ \text{div}\ \textbf{grad}\ f + \textbf{grad}\ g\cdot\textbf{grad}\ f)\ d\Omega = \int_{\partial\Omega} g\ \textbf{grad}\ f\cdot\hat{\boldsymbol{n}}\, d\partial\Omega \tag{B.17}$$

or

$$\int_\Omega g\ \text{div}\ \textbf{grad}\ f\, d\Omega = \int_{\partial\Omega} g\textbf{grad}\ f\cdot\hat{\boldsymbol{n}}\, d\partial\Omega - \int_\Omega \textbf{grad}\ g\cdot\textbf{grad}\ f\, d\Omega. \tag{B.18}$$

Eq.(B.18) is sometimes referred to as *integration by parts*.

B.4 Weak formulations

B.4.1 Volume problem

Consider the volume problem Eqs(2.4), *i.e.*

$$\begin{cases} \left(-\text{div}\ \textbf{grad}\ - k^2\right) w = f & \text{in } \Omega_i, & \text{(B.19)}\\ \left(\hat{\boldsymbol{n}}\cdot\textbf{grad}\ + \mathcal{B}\right) w = u_\infty & \text{on } \partial\Omega_i^\infty, & \text{(B.20)}\\ w = u_d & \text{on } \partial\Omega_i^d, & \text{(B.21)}\\ \hat{\boldsymbol{n}}\cdot\textbf{grad}\ w = u_n & \text{on } \partial\Omega_i^n, & \text{(B.22)}\\ \mathcal{J}_i w = s_{im} & \text{on } \Gamma_{im},\ \forall m\in D_i. & \text{(B.23)} \end{cases}$$

Volume partial differential equation Eq.(B.19) (T1) Multiplying Eq.(B.19) by a test function $\tilde{w}$ and then integrating over Ω_i yields

$$-\int_{\Omega_i} \text{div}\ \textbf{grad}\ w\,\tilde{w}\, d\Omega_i - \omega^2\int_{\Omega_i}\left(\frac{k}{\omega}\right)^2 w\tilde{w}\, d\Omega_i - \int_{\Omega_i} f\tilde{w}\, d\Omega_i = 0 \tag{B.24}$$

which reduces to

$$\text{(T1)}\quad -\left(\underbrace{\int_{\partial\Omega_i}\tilde{w}\ \textbf{grad}\ w\cdot\hat{\boldsymbol{n}}\, d\partial\Omega_i}_{\text{A}} - \int_{\Omega_i}\textbf{grad}\ \tilde{w}\cdot\textbf{grad}\ w\, d\Omega_i\right) - \omega^2\int_{\Omega_i}\left(\frac{k}{\omega}\right)^2\tilde{w}\, d\Omega_i - \int_{\Omega_i} f\tilde{w}\, d\Omega_i = 0 \tag{B.25}$$

using integration by parts Eq.(B.18).

Absorbing boundary condition Eq.(B.20) (T2) Multiplying Eq.(B.20) by the test function $\tilde{w}$ and then integrating over $\partial\Omega_i^\infty$ yields

$$\text{(T2)} \quad \underbrace{\int_{\partial\Omega_i^\infty} \tilde{w}\ \mathbf{grad}\ w \cdot \hat{\boldsymbol{n}}\, d\partial\Omega_i^\infty}_{\text{B}} + \int_{\partial\Omega_i^\infty} \tilde{w}\mathcal{B}w\, d\partial\Omega_i^\infty - \int_{\partial\Omega_i^\infty} \tilde{w}u_\infty\, d\partial\Omega_i^\infty = 0. \tag{B.26}$$

Dirichlet boundary condition Eq.(B.21) (T3) The Dirichlet boundary condition Eq.(B.21) is automatically verified because $\tilde{w}$ cancels on $\partial\Omega_i^d$ and thus

$$\text{(T3)} \quad \underbrace{\int_{\partial\Omega_i^d} \tilde{w}\ \mathbf{grad}\ w \cdot \hat{\boldsymbol{n}}\, d\partial\Omega_i^d}_{\text{C}} = 0. \tag{B.27}$$

Neumann boundary condition Eq.(B.22) (T4) Multiplying Eq.(B.22) by a test function $\tilde{w}$ and then integrating over $\partial\Omega_i^n$ yields

$$\text{(T4)} \quad \underbrace{\int_{\partial\Omega_i^n} \tilde{w}\ \mathbf{grad}\ w \cdot \hat{\boldsymbol{n}}\, d\partial\Omega_i^n}_{\text{D}} - \int_{\partial\Omega_i^n} \tilde{w}u_n\, d\partial\Omega_i^\infty = 0. \tag{B.28}$$

Transmission boundary conditions Eq.(B.23) (T5) Multiplying each ($\forall m \in D_i$) of Eqs(B.23) by a test function $\tilde{w}$ and then integrating over $\partial\Omega_i^\infty$ yields

$$\text{(T5)} \quad \underbrace{\int_{\Gamma_{im}} \tilde{w}\ \mathbf{grad}\ w \cdot \hat{\boldsymbol{n}}\, d\Gamma_{im}}_{\text{E}} + \int_{\Gamma_{im}} \tilde{w}\mathcal{S}w\, d\Gamma_{im} - \int_{\Gamma_{im}} \tilde{w}s_{im}\, d\Gamma_{im} = 0, \quad \forall m \in D_i. \tag{B.29}$$

By definition, boundaries satisfy

$$\partial\Omega_i = \partial\Omega_i^\infty \cup \partial\Omega_i^d \cup \partial\Omega_i^n \cup_{m\in D_i} \Gamma_{im} \tag{B.30}$$

such that

$$\text{A} = \text{B} + \text{C} + \text{D} + \text{E}. \tag{B.31}$$

Summing terms T1 to T5 then gives

$$\begin{aligned} &\int_{\Omega_i} \mathbf{grad}\ \tilde{w} \cdot \mathbf{grad}\ w\, d\Omega_i - \omega^2 \int_{\Omega_i} \left(\frac{k}{\omega}\right)^2 \tilde{w}w\, d\Omega_i \\ &\qquad + \int_{\partial\Omega_i^\infty} \tilde{w}\mathcal{B}w\, d\partial\Omega_i^\infty + \sum_{m\in D_i} \int_{\Gamma_{im}} \tilde{w}\mathcal{S}w\, d\Gamma_{im} \\ &= \int_{\Omega_i} \tilde{w}f\, d\Omega_i + \int_{\partial\Omega_i^\infty} \tilde{w}u_\infty\, d\partial\Omega_i^\infty + \int_{\partial\Omega_i^n} \tilde{w}u_n\, d\partial\Omega_i^\infty + \sum_{m\in D_i} \int_{\Gamma_{im}} \tilde{w}s_{im}\, d\Gamma_{im}. \end{aligned} \tag{B.32}$$

B.4.2 Surface equation

Consider the surface equation Eq.(2.16), *i.e.*

$$\mathcal{A}_{ji}(s_{im}\,\forall m \in D_i) = -s_{ij} + (\mathcal{S}_i + \mathcal{S}_j)\, w. \tag{B.33}$$

Multiplying Eq.(B.33) by a test function $\tilde{\mathcal{A}}_{ji}$ and then integrating over the common boundary Γ_{ij} yields

$$\boxed{\int_{\Gamma_{ij}} \tilde{\mathcal{A}}_{ji}\mathcal{A}_{ji}\, d\Gamma_{ij} = -\int_{\Gamma_{ij}} \tilde{\mathcal{A}}_{ji}s_{ij}\, d\Gamma_{ij} + \int_{\Gamma_{ij}} \tilde{\mathcal{A}}_{ji}\,(\mathcal{S}_i + \mathcal{S}_j)\, w\, d\Gamma_{ij}.} \tag{B.34}$$

B.4.3 Boundary operators

In Section 1.6 and 2.4, particular forms for the operators $\mathcal{B}$ and $\mathcal{S}$ are derived. All these operators express as

$$\mathcal{K} = \mathcal{B} \text{ or } \mathcal{S} = a - b\text{div}_\Sigma\, \mathbf{grad}_\Sigma \tag{B.35}$$

where the expression of a and b depend on the considered operator and can be found in sections 1.6 and 2.4.
The weak form term

$$\int_\Sigma \tilde{v}\mathcal{K}v\, d\Sigma \tag{B.36}$$

where $v = w$ or $\mathcal{A}_{ji}$, $\tilde{v} = \tilde{w}$ or $\tilde{\mathcal{A}}_{ji}$ and $\Sigma = \partial\Omega_i^\infty$ or Γ_{ji}, then becomes

$$\boxed{\int_\Sigma \tilde{v}\mathcal{K}v\, d\Sigma = a\int_\Sigma v\tilde{v}\, d\Sigma - b\left[\tilde{v}\,\hat{\boldsymbol{t}}\cdot\mathbf{grad}_\Sigma v\right]_{\partial\Sigma} + b\int_\Sigma \mathbf{grad}_\Sigma\, \tilde{v}\cdot\mathbf{grad}_\Sigma\, v\, d\Sigma} \tag{B.37}$$

where $\hat{\boldsymbol{t}}$ is the tangent to Σ.

B.5 Mass and stiffness matrices definitions

$$\left[\underline{\underline{K}}^{vv}_{\Omega_i}\right]_{pq} \triangleq \int_{\Omega_i} \mathbf{grad}\,\psi_p^v \cdot \mathbf{grad}\,\psi_q^v\, d\Omega_i, \qquad \forall p \in N_{\Omega_i}, q \in N_{\Omega_i} \tag{B.38}$$

$$\left[\underline{\underline{M}}^{vv}_{\Omega_i}\right]_{pq} \triangleq \int_{\Omega_i} \psi_p^v\psi_q^v\, d\Omega_i, \qquad \forall p \in N_{\Omega_i}, q \in N_{\Omega_i} \tag{B.39}$$

$$\left[\underline{\underline{K}}^{vv}_{\partial\Omega_i^\infty}\right]_{pq} \triangleq \int_{\partial\Omega_i^\infty} \mathbf{grad}_{\partial\Omega_i^\infty}\psi_p^v \cdot \mathbf{grad}_{\partial\Omega_i^\infty}\psi_q^v\, d\partial\Omega_i^\infty, \forall p \in N_{\Omega_i}, q \in N_{\Omega_i} \tag{B.40}$$

$$\left[\underline{\underline{M}}^{vv}_{\partial\Omega_i^\infty}\right]_{pq} \triangleq \int_{\partial\Omega_i^\infty} \psi_p^v\psi_q^v\, d\partial\Omega_i^\infty, \qquad \forall p \in N_{\Omega_i}, q \in N_{\Omega_i} \tag{B.41}$$

$$\left[\underline{\underline{K}}^{vv}_{\Gamma_{im}}\right]_{pq} \triangleq \int_{\Gamma_{im}} \mathbf{grad}_{\Gamma_{im}}\psi_p^v \cdot \mathbf{grad}_{\Gamma_{im}}\psi_q^v\, d\Gamma_{im}, \qquad \forall p \in N_{\Omega_i}, q \in N_{\Omega_i} \tag{B.42}$$

$$\left[\underline{\underline{M}}^{vv}_{\Gamma_{im}}\right]_{pq} \triangleq \int_{\Gamma_{im}} \psi_p^v\psi_q^v\, d\Gamma_{im}, \qquad \forall p \in N_{\Omega_i}, q \in N_{\Omega_i} \tag{B.43}$$

$$\left[\underline{f}_{\Omega_i}\right]_p \triangleq \int_{\Omega_i} f\psi_p^v\, d\Omega_i, \qquad \forall p \in N_{\Omega_i} \tag{B.44}$$

$$\left[\underline{u}_{\infty,\partial\Omega_i}\right]_p \triangleq \int_{\partial\Omega_i^a} u_\infty \psi_p^v \, d\partial\Omega_i^a, \qquad \forall p \in N_{\Omega_i} \tag{B.45}$$

$$\left[\underline{u}_{n,\partial\Omega_i}\right]_p \triangleq \int_{\partial\Omega_i^n} u_n \psi_p^v \, d\partial\Omega_i^n, \qquad \forall p \in N_{\Omega_i} \tag{B.46}$$

$$\left[\underline{\underline{K}}^{sv}_{\Gamma_{ij}}\right]_{pq} \triangleq \int_{\Gamma_{ij}} \mathbf{grad}_{\Gamma_{im}} \, \psi_p^s \cdot \mathbf{grad}_{\Gamma_{im}} \, \psi_q^v \, d\Gamma_{ij}, \qquad \forall p \in N_{\Gamma_{ij}}, q \in N_{\Omega_i} \tag{B.47}$$

$$\left[\underline{\underline{M}}^{vs}_{\Gamma_{ij}}\right]_{pq} \triangleq \int_{\Gamma_{ij}} \psi_p^v \psi_q^s \, d\Gamma_{ij}, \qquad \forall p \in N_{\Omega_i}, q \in N_{\Gamma_{ij}} \tag{B.48}$$

$$\left[\underline{\underline{M}}^{ss}_{\Gamma_{ij}}\right]_{pq} \triangleq \int_{\Gamma_{ij}} \psi_p^s \psi_q^s \, d\Gamma_{ij}, \qquad \forall p \in N_{\Gamma_{ij}}, q \in N_{\Gamma_{ij}}. \tag{B.49}$$

B.6 Link between time and frequency domains for time harmonic fields

Time harmonic fields at pulsation ω_0 are given in the frequency domain by

$$\begin{aligned} \boldsymbol{f}(\omega) &= \frac{1}{2}\left(\boldsymbol{f}(\omega_0)\delta(\omega-\omega_0) + \boldsymbol{f}(-\omega_0)\delta(\omega+\omega_0)\right) \\ &= \frac{1}{2}\left(\boldsymbol{f}(\omega_0)\delta(\omega-\omega_0) + \boldsymbol{f}(\omega_0)^*\delta(\omega+\omega_0)\right) \end{aligned} \tag{B.50}$$

and in the time domain by

$$\mathbf{f}(t) = \mathrm{Re}\left(\boldsymbol{f}(\omega_0)\right)\cos\left(\omega_0 t\right) - \mathrm{Im}\left(\boldsymbol{f}(\omega_0)\right)\sin\left(\omega_0 t\right) \tag{B.51}$$

$$\triangleq \boldsymbol{f}_r \cos\left(\omega_0 t\right) - \boldsymbol{f}_i \sin\left(\omega_0 t\right). \tag{B.52}$$

The scalar product of two such fields is given by

$$\begin{aligned} \mathbf{f}\cdot\mathbf{g} &= \boldsymbol{f}_r\cdot\boldsymbol{g}_r\cos^2\left(\omega_0 t\right) + \boldsymbol{f}_i\cdot\boldsymbol{g}_i\sin^2\left(\omega_0 t\right) - \boldsymbol{f}_r\cdot\boldsymbol{g}_i\cos\left(\omega_0 t\right)\sin\left(\omega_0 t\right) - \boldsymbol{f}_i\cdot\boldsymbol{g}_r\cos\left(\omega_0 t\right)\sin\left(\omega_0 t\right) \\ &= \frac{1}{2}\left[\boldsymbol{f}_r\cdot\boldsymbol{g}_r + \boldsymbol{f}_i\cdot\boldsymbol{g}_i + \boldsymbol{f}_r\cdot\boldsymbol{g}_r\cos\left(2\omega_0 t\right) - \boldsymbol{f}_i\cdot\boldsymbol{g}_i\cos\left(2\omega_0 t\right) - \boldsymbol{f}_r\cdot\boldsymbol{g}_i\sin\left(2\omega_0 t\right) - \boldsymbol{f}_i\cdot\boldsymbol{g}_r\sin\left(2\omega_0 t\right)\right]. \end{aligned} \tag{B.53}$$

The time average over one period $T = \frac{2\pi}{\omega_0}$ is given by

$$\begin{aligned} \frac{1}{T}\int_0^T \boldsymbol{f}\cdot\boldsymbol{g}\, dt &= \frac{1}{2}\left[\boldsymbol{f}_r\cdot\boldsymbol{g}_r + \boldsymbol{f}_i\cdot\boldsymbol{g}_i\right] \\ &= \frac{1}{2}\mathrm{Re}\left(\boldsymbol{f}\cdot\boldsymbol{g}^*\right). \end{aligned} \tag{B.54}$$

Using similar arguments also yields

$$\begin{aligned} \frac{1}{T}\int_0^T \mathbf{f}\times\mathbf{g}\, dt &= \frac{1}{2}\left[\boldsymbol{f}_r\times\boldsymbol{g}_r + \boldsymbol{f}_i\times\boldsymbol{g}_i\right] \\ &= \frac{1}{2}\mathrm{Re}\left(\boldsymbol{f}\times\boldsymbol{g}^*\right). \end{aligned} \tag{B.55}$$

B.7 Poynting's theorem

B.7.1 Time domain

Consider the time domain Poynting's vector $\mathbf{s}$ defined as

$$\boldsymbol{s} = \mathbf{e} \times \mathbf{h} \tag{B.56}$$

and an arbitrary closed volume V whose surface is denoted by $S\,(= S_{\text{in}} \bigcup S_{\text{ext}})$ as represented in Figure B.3.

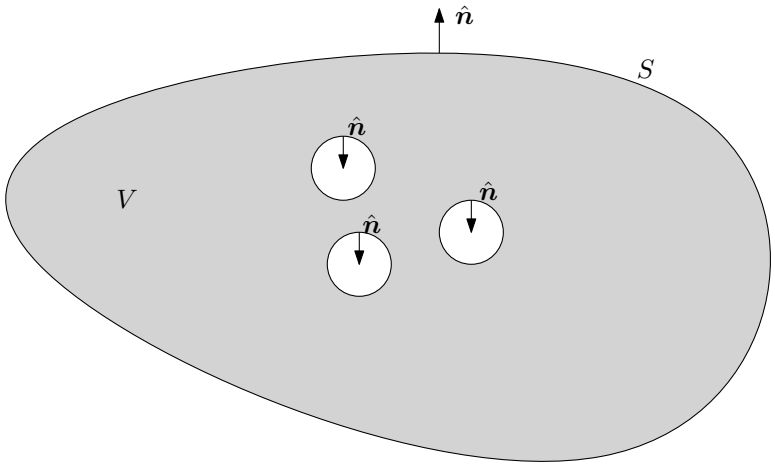


Figure B.3: Arbitrary volume on which the Poynting's theorem is applied.

Using the divergence theorem Eq.(B.16), one has

$$\int_S \mathbf{s} \cdot \hat{\boldsymbol{n}}\, dS = \int_V \text{div}\, \mathbf{s}\, dV. \tag{B.57}$$

The divergence term can then be developed using successively the definition of the Poynting's vector Eq.(B.56), the Ampère-Maxwell's Eq.(1.1) and the Faraday's law Eq.(1.2) as[2]

$$\begin{aligned} \text{div}\, \mathbf{s} &= (\textbf{curl}\, \mathbf{e}) \cdot \mathbf{h} - \mathbf{e} \cdot (\textbf{curl}\, \mathbf{h}) \\ &= -\mu \mathbf{h} \cdot \frac{\partial \mathbf{h}}{\partial t} - \boldsymbol{e} \cdot \left(\epsilon \frac{\partial \mathbf{e}}{\partial t} + \mathbf{j}_c + \mathbf{j}_a \right) \\ &= -\frac{1}{2} \mu \frac{\partial \left\| \mathbf{h} \right\|^2}{\partial t} - \frac{1}{2} \epsilon \frac{\partial \left\| \mathbf{e} \right\|^2}{\partial t} - \sigma \left\| \mathbf{e} \right\|^2 - \mathbf{e} \cdot \mathbf{j}_a. \end{aligned} \tag{B.58}$$

Then inserting Eq.(B.58) in the divergence theorem yields

$$\boxed{-\int_V \mathbf{e} \cdot \mathbf{j}_a\, dV = \frac{\partial}{\partial t} \int_V \frac{1}{2} \left(\mu \left\| \mathbf{h} \right\|^2 + \epsilon \left\| \mathbf{e} \right\|^2 \right) dV + \int_V \sigma \left\| \mathbf{e} \right\|^2 dV + \int_S \mathbf{s} \cdot \hat{\boldsymbol{n}}\, dS.} \tag{B.59}$$

[2]Isotropic linear constitutive laws are considered, *i.e.* $\mathbf{e} = \epsilon\,(\boldsymbol{x})\, \mathbf{d}$ and $\mathbf{b} = \mu\,(\boldsymbol{x})\, \mathbf{h}$.

The physical interpretation of the terms appearing in the power conservation Eq.(B.59) is as folows:

- $$\mathrm{P_s} \triangleq -\int_V \mathbf{e}\cdot\mathbf{j}_a\, dV \tag{B.60}$$
  is the *power injected* by sources. It is positive when $\mathbf{e}$ and $\mathbf{j}_a$ have opposite direction what makes sense.

- $$\mathrm{S} \triangleq \frac{\partial}{\partial t}\int_V \left(\mu\left\|\mathbf{h}\right\|^2 + \epsilon\left\|\mathbf{e}\right\|^2\right)\, dV \tag{B.61}$$
  is the time variation of the *stored electromagnetic energy* inside volume V.

- $$\mathrm{L} \triangleq \int_V \sigma\left\|\mathbf{e}\right\|^2\, dV \tag{B.62}$$
  are the *losses* generated by conduction. As $\sigma \geq 0$, this term is logically always positive.

- $$\mathrm{R} \triangleq \int_S \mathbf{s}\cdot\hat{\boldsymbol{n}}\, dS \tag{B.63}$$
  is the power radiated through the surface S.

B.7.2 Frequency domain

Consider the Poynting's vector $\boldsymbol{s}$ defined as

$$\boldsymbol{s} = \frac{1}{2}\boldsymbol{e}\times\boldsymbol{h}^*. \tag{B.64}$$

Using the link between time and frequency domain (*cfr* Appendix B.6)), the relation

$$\frac{1}{T}\int_0^T \mathbf{s}\, dt = \mathrm{Re}\left(\boldsymbol{s}\right) \tag{B.65}$$

holds between time and frequency domain for time harmonic fields.
Once again using the divergence theorem, one has

$$\int_S \boldsymbol{s}\cdot\hat{\boldsymbol{n}}\, dS = \int_V \mathrm{div}\ \boldsymbol{s}\, dV. \tag{B.66}$$

The divergence can then be developed using successively the definition of the Poynting's vector Eq.(B.64), the Ampère-Maxwell's Eq.(1.7) and the Faraday's law Eq.(1.8) as

$$\begin{aligned} 2\,\mathrm{div}\ \boldsymbol{s} &= \left(\mathbf{curl}\ \boldsymbol{e}\right)\cdot h^* - \boldsymbol{e}\cdot\left(\mathbf{curl}\ \boldsymbol{h}^*\right) \\ &= -i\omega\mu_0\boldsymbol{h}\cdot\boldsymbol{h}^* - \boldsymbol{e}\cdot\left(i\omega\tilde{\epsilon}\boldsymbol{e} + \boldsymbol{j}_a\right)^* \\ &= -i\omega\mu_0\left\|\boldsymbol{h}\right\|^2 + i\omega\mathrm{Re}\left(\tilde{\epsilon}\right)\left\|\boldsymbol{e}\right\|^2 + \omega\mathrm{Im}\left(\tilde{\epsilon}\left\|\boldsymbol{e}\right\|^2\right) - \boldsymbol{e}\cdot\boldsymbol{j}_a{}^*. \end{aligned} \tag{B.67}$$

Then inserting Eq.(B.67) in the divergence theorem yields

$$\boxed{-\frac{1}{2}\int_V \boldsymbol{e}\cdot\boldsymbol{j}_a{}^*\,dV = i\omega\frac{1}{2}\int_V \left(\mu_0\left\|\boldsymbol{h}\right\|^2 - \mathrm{Re}\left(\tilde{\epsilon}\right)\left\|\boldsymbol{e}\right\|^2\right)\,dV + \frac{1}{2}\int_V \tilde{\sigma}\left\|\boldsymbol{e}\right\|^2\,dV + \int_S \boldsymbol{s}\cdot\hat{\boldsymbol{n}}\,dS.} \tag{B.68}$$

Taking the real part of the complex Poynting theorem Eq.(B.68), a *time average power conservation theorem* is obtained. Indeed the real part is given by

$$\boxed{-\frac{1}{2}\int_V \mathrm{Re}\left(\boldsymbol{e}\cdot\boldsymbol{j}_a{}^*\right)\,dV = \frac{1}{2}\int_V \tilde{\sigma}\left\|\boldsymbol{e}\right\|^2\,dV + \int_S \mathrm{Re}\left(\boldsymbol{s}\right)\cdot\hat{\boldsymbol{n}}\,dS.} \tag{B.69}$$

The physical interpretation of the terms appearing in the power conservation Eq.(B.69) is deduced from time domain and is as folows:

- $$P_s \triangleq -\frac{1}{2}\int_V \mathrm{Re}\left(\boldsymbol{e}\cdot\boldsymbol{j}_a{}^*\right)\,dV = \frac{1}{T}\int_0^T \mathrm{P_s}\,dt \tag{B.70}$$
  is the power injected by the sources.

- $$R \triangleq \int_S \mathrm{Re}\left(\boldsymbol{s}\right)\cdot\hat{\boldsymbol{n}}\,dS = \frac{1}{T}\int_0^T \mathrm{R}\,dt \tag{B.71}$$
  is the power radiated through the surface S.

- $$L \triangleq \frac{1}{2}\int_V \tilde{\sigma}\left\|\boldsymbol{e}\right\|^2\,dV = \frac{1}{T}\int_0^T \mathrm{L}\,dt \tag{B.72}$$
  are the loss generated by the equivalent conductivity.

A similar principle can be obtained for the reactive power through the imaginary part of Eq.(B.68).

Transverse electric (TE)

If the electric field $\boldsymbol{e}$ is transverse, *i.e.*

$$\boldsymbol{e} = u\hat{\boldsymbol{z}} \tag{B.73}$$

then in regions where there are no sources (*i.e.* $\boldsymbol{j}_a = \boldsymbol{0}$), the Poynting's vector expresses as

$$\begin{aligned}\boldsymbol{s} &\triangleq \frac{1}{2}\boldsymbol{e}\times\boldsymbol{h}^*\\ &= \frac{1}{2}u\hat{\boldsymbol{z}}\times\frac{\mathbf{curl}\,\boldsymbol{e}^*}{i\omega\mu_0}\\ &= \frac{1}{2}\frac{u}{i\omega\mu_0}\hat{\boldsymbol{z}}\times\mathbf{curl}\,\boldsymbol{e}^*\\ &= \frac{1}{2}\frac{1}{i\omega\mu_0}u\mathbf{grad}\,u^*.\end{aligned} \tag{B.74}$$

The different terms of the power conservation Eq.(B.69) then take the following forms:

$$P_s \triangleq -\frac{1}{2}\int_V \text{Re}\left(u j_a^*\right)\, dV, \tag{B.75}$$

$$R \triangleq \frac{1}{2}\frac{1}{\omega\mu_0}\int_S \text{Im}\left(u\,\textbf{grad}\, u^*\right)\, dS, \tag{B.76}$$

$$L \triangleq \frac{1}{2}\int_V \tilde{\sigma}\left\|u\right\|^2\, dV. \tag{B.77}$$

B.8 Duality between losses minimization and radiation maximization

Consider three symmetric matrices denoted by $\underline{\underline{P_s}}$, $\underline{\underline{L}}$ and $\underline{\underline{R}}$ such that

$$\underline{j}\left(\underline{\underline{P_s}} - \underline{\underline{L}} - \underline{\underline{R}}\right)\underline{j} = 0, \quad \forall \underline{j} \tag{B.78}$$

which is equivalent to consider that $\underline{\underline{P_s}} - \underline{\underline{L}} - \underline{\underline{R}}$ is skew symmetric. Taking the gradient w.r.t. $\underline{j}$ of Eq.(B.78) yields

$$\left(\underline{\underline{P_s}} - \underline{\underline{L}} - \underline{\underline{R}}\right)\underline{j} = 0 \tag{B.79}$$

as the three matrices are symmetric. Then introducing $\lambda^R \in \mathbb{R}$, the preceding equation can be written

$$\left(\lambda^R\underline{\underline{P_s}} - \underline{\underline{R}}\right)\underline{j} + \left((1-\lambda^R)\underline{\underline{P_s}} - \underline{\underline{L}}\right)\underline{j} = 0. \tag{B.80}$$

Therefore if a $\{\lambda_n^R, \underline{j_n}\}$ couple verifies

$$\left(\lambda_n^R\underline{\underline{P_s}} - \underline{\underline{R}}\right)\underline{j_n} = 0 \tag{B.81}$$

then

$$\left((1-\lambda_n^R)\underline{\underline{P_s}} - \underline{\underline{L}}\right)\underline{j_n} = 0. \tag{B.82}$$

B.9 Link between electric losses and thermal elevation

The heat equation using Fourier conduction law is given by

$$\rho c\frac{\partial T}{\partial t} - \text{ div }\left(k\,\textbf{grad}\, T\right) = \dot{q} \tag{B.83}$$

where k [W/m/K] is the *thermal conductivity* and $\dot{q}$ [W/m^3] are the *volume heat sources*.
In the particular case of human tissues heated by Joules effect (no polarization losses are considered here, *i.e.* $\text{Im}\left(\epsilon\right) = 0$), then the source term is

$$\mathbf{e}\cdot\mathbf{j} = \sigma\left\|\mathbf{e}(t)\right\|^2. \tag{B.84}$$

For time-harmonic fields at pulsation ω_0 (*cfr* Eq.(B.53)) the dissipation is

$$\sigma\left\|\mathbf{e}\right\|^2 = \quad = \frac{1}{2}\sigma\left[\left\|\boldsymbol{e}_r\right\|^2 + \left\|\boldsymbol{e}_i\right\|^2 + \left\|\boldsymbol{e}_r\right\|^2\cos\left(2\omega_0 t\right) - \left\|\boldsymbol{e}_i\right\|^2\cos\left(2\omega_0 t\right) - 2\boldsymbol{e}_r\cdot\boldsymbol{e}_i\sin\left(2\omega_0 t\right)\right]. \tag{B.85}$$

If the temperature is initially homogeneous inside the media and equal to T_0 then initially the spatial gradient and thus the conduction term cancels such that heat equation Eq.(B.83) becomes

$$\rho c \frac{\partial T}{\partial t} = \sigma \left\| \mathbf{e} \right\|^2 (t) \tag{B.86}$$

and the solution is

$$\begin{aligned} T(t) &= T_0 + \frac{1}{c} \frac{\sigma \left\| \boldsymbol{e}(\omega_0) \right\|^2}{2\rho} \left(t + \mathcal{O}(\omega^{-1}) \right) \\ &\approx T_0 + \frac{1}{c} \frac{\sigma \left\| \boldsymbol{e}(\omega_0) \right\|^2}{2\rho} t \\ &\approx T_0 + \frac{1}{c} \mathrm{SAR}\, t. \end{aligned} \tag{B.87}$$

The oscillation terms $\mathcal{O}(\omega_0^{-1})$ with periodicity $2\pi/\omega_0$ are negligible in the GigaHertz range as $2\pi/\omega_0 \approx 1$ [ns].
Now the solution is known, the initial assumption can be checked, *i.e.*

$$\begin{aligned} & \frac{k \,\mathrm{div}\, (\mathbf{grad}\, T)}{\frac{1}{2}\sigma \left\| \boldsymbol{e}(\omega_0) \right\|^2} && \ll 1 \\ \Rightarrow\ & \frac{kt}{\rho c} \frac{\mathrm{div}\, \left(\mathbf{grad}\, \left\| \boldsymbol{e} \right\|^2 \right)}{\left\| \boldsymbol{e} \right\|^2} && \ll 1 \\ \Rightarrow\ & \frac{kt}{\rho c \lambda^2} && \ll 1 \\ \Rightarrow\ & t && \ll \tau \triangleq \frac{\rho c \lambda^2}{k}. \end{aligned} \tag{B.88}$$

Using typical thermal [22] and electrical [10] data from muscles yields[3] $\frac{k}{\rho c} \approx 10^{-7}$ [m^2/s] and $\lambda \approx 1$ [cm] at $f = 3$ [GHz] such that

$$\tau \approx 17 \text{ [min]}. \tag{B.89}$$

After this time τ, the conduction term is no longer negligible and the entire partial differential equation Eq.(B.83) must be solved. In the particular case of biological tissues, the classical heat equation Eq.(B.83) is often substituted by the bio-heat equation [14] which is designed to take into account heat convection through the blood, among others.

[3] $\rho = 1.2$ [g/cm^3], $c = 0.9$ [cal/g/K], $k = 0.001$ [cal/cm/s/K].

Appendix C

Additional tables and figures

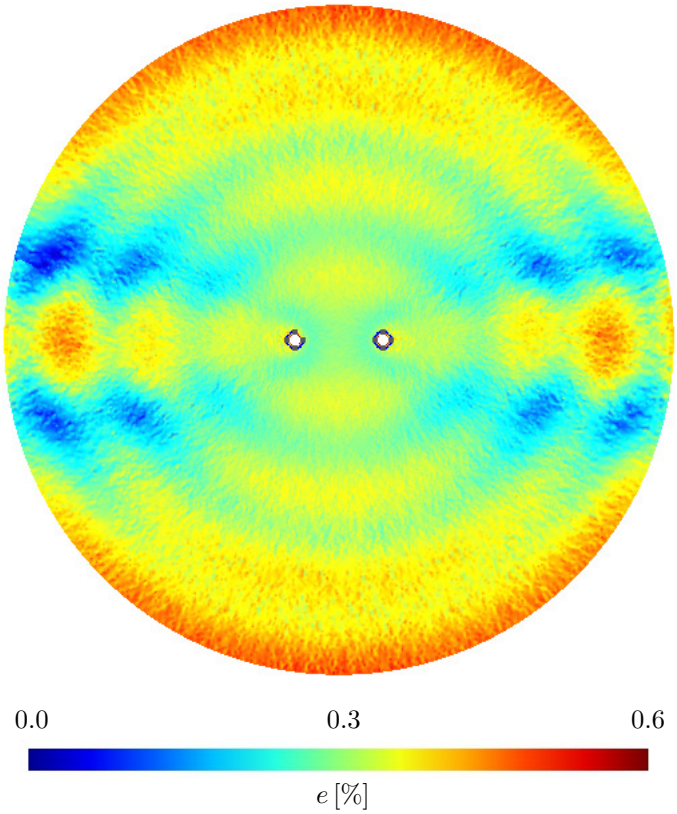


Figure C.1: Map of the error e for Model Problem III. Krylov solver and OO2 transmission conditions are used. The tolerance is set to $\epsilon_{\text{it}} = 10^{-5}$.

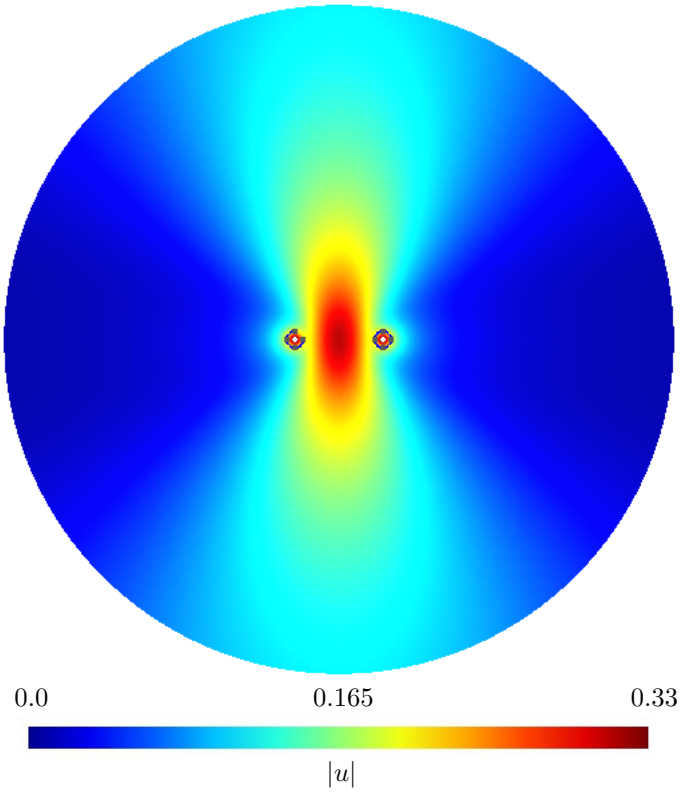


Figure C.2: Norm of the electric field u for Model Problem III consisting in a 2-source array.

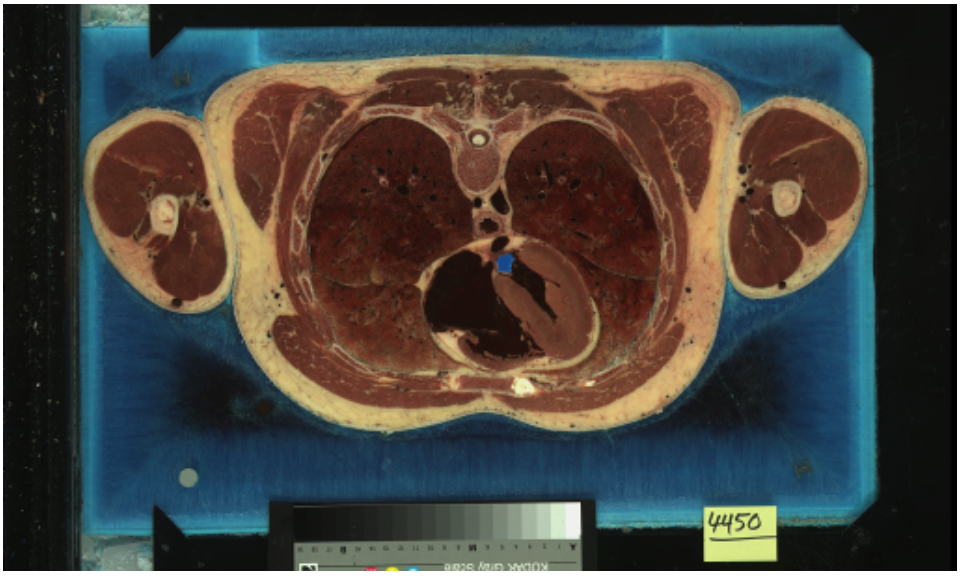


Figure C.3: Color axial anatomical images from the Visible Human Project sponsored by the USA National Library of Medicine [29] used to create the geometrical human body model of Chapter 8.

	Model Problem				Application		Units
	I	II	III	IV			
k_0	1	1	1	-	-	-	[1/m]
a	0.6	-	-	-	-	-	[m]
b	{6,30}	12	12	0.3	0.3	0.3	[m]
x_p	{0,-3}	-	-	-	-	-	[m]
ABC	{CT0, CT2}	CT2	CT2	CT2	CT2	CT2	
N_d	1	1	5	5	6	6	[-]
h_0	0.13	0.13	0.13	0.0033	0.025	0.008	[m]
ϵ	-	{0.3,0.6}	0.3	0.005	0.003	5×10^{-4}	[m]
m	-	{5,17}	9	5	3	3	[-]
Source	-	{CST, GSN}	GSN	GSN	CST	CST	
TC	-	-	{T0BC, T2BC, OO2BC}	OO2BC	OO2BC	OO2BC	
Solver	-	-	{Jacobi, Krylov}	Krylov	Krylov	Krylov	
d	-	-	π	0.05	-	0.005	[m]
ϵ_{it}	-	-	$[10^{-1} : 10^{-5}]$	10^{-6}	10^{-6}	10^{-6}	[-]
f	-	-	-	3×10^9	10^8	3×10^9	[Hz]
λ_0	-	-	-	0.1	3	0.1	[m]
λ_m	-	-	-	0.01	0.3	0.01	[m]
ρ_m	-	-	-	0.05	-	-	[m]
d_m	-	-	-	0.1	-	-	[m]
h_m	-	-	-	0.001	0.005	8×10^{-4}	[m]

Table C.1: Parameters related to the elementary solutions for the numerical simulations of Part III.

		Bone (cortical)	Fat (not infiltrated)	Heart	Lung (inflated)	Muscle
ϵ_∞	[-]	2.5	2.5	4.0	2.5	4.0
$\Delta\epsilon_1$	[-]	10.0	3.0	50.0	18.0	50.0
τ_1	[ps]	13.26	7.96	7.96	7.96	7.23
α_1	[-]	0.20	0.20	0.10	0.10	0.10
$\Delta\epsilon_2$	[-]	180	15	1200	500	7000
τ_2	[ns]	79.58	15.92	159.15	63.66	353.68
α_2	[-]	0.20	0.10	0.05	0.10	0.10
$\Delta\epsilon_3$	[-]	5.0×10^3	3.3×10^4	4.5×10^5	2.5×10^5	1.2×10^6
τ_3	[μs]	159.15	159.15	72.34	159.15	318.31
α_3	[-]	0.20	0.05	0.22	0.20	0.10
$\Delta\epsilon_4$	[-]	1.0×10^5	1.0×10^7	2.5×10^7	4.0×10^7	2.5×10^7
τ_4	[ms]	15.915	7.958	4.547	7.958	2.274
α_4	[-]	0.00	0.01	0.00	0.00	0.00
σ	[S]	0.0200	0.0100	0.0500	0.0300	0.2000

Table C.2: Cole Cole dispersions and conductivity parameters for several biological tissues considered in the human body model [10].

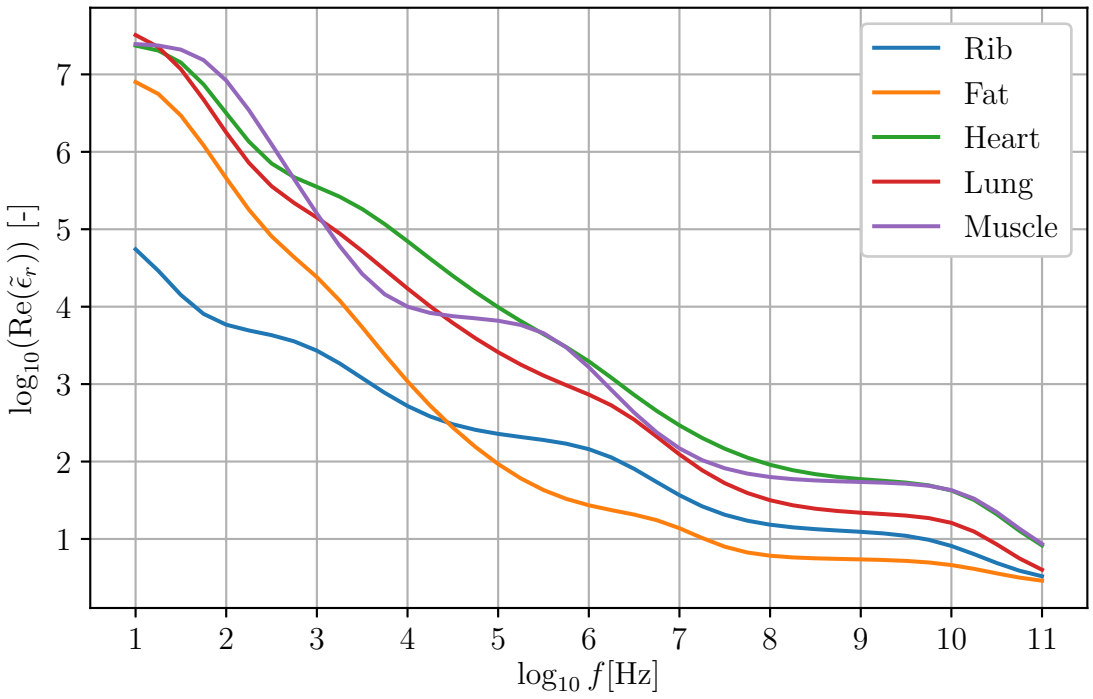


Figure C.4: Real part of the permittivity $\mathrm{Re}\,(\tilde{\epsilon}_r)$ for the different tissues considered in the human body model.

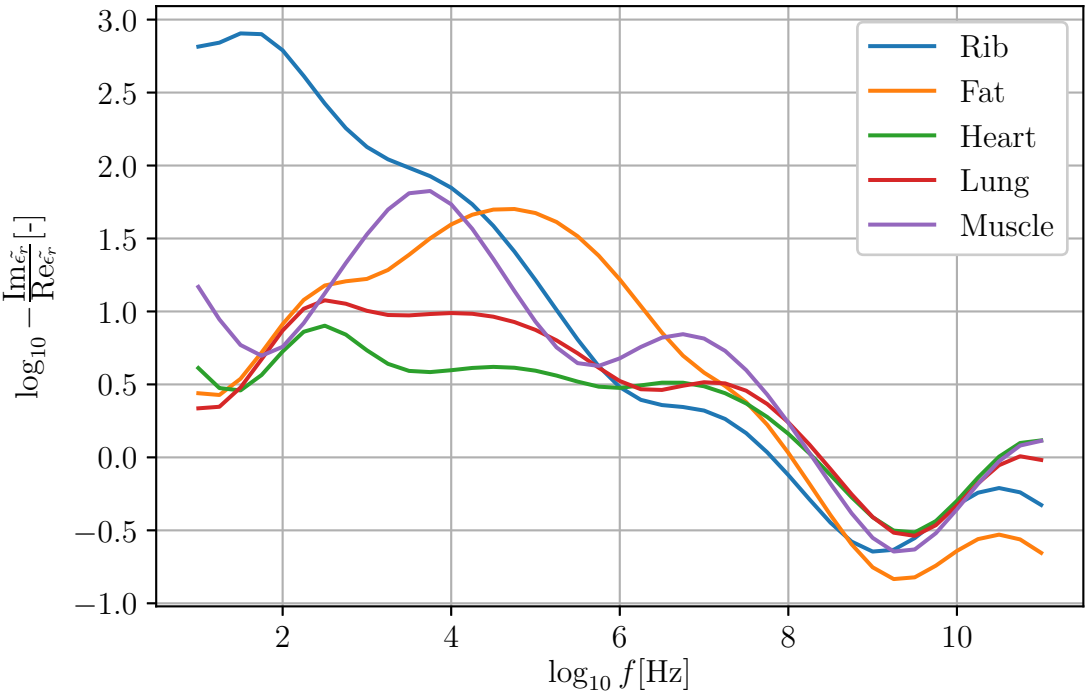


Figure C.5: Conductor ratio $-\frac{\mathrm{Im}(\tilde{\epsilon}_r)}{\mathrm{Re}(\tilde{\epsilon}_r)} = \frac{\tilde{\sigma}}{\omega\epsilon_0 \mathrm{Re}(\tilde{\epsilon}_r)}$ for the different tissues considered in the human body model.

	$N_p = 1$	$N_p = 2$	$N_p = 5$	$N_p = 9$
j_1	(52.8;0)	(39.8;-49.6)	(35;-46)	(28.5;-72.1)
j_2		(50.7;180)	(44.2;167.1)	(37;144.3)
j_3			(17.1;-56.7)	(29.3;19.44)
j_4			(5.7;91.2)	(8.1;41.4)
j_5			(23.6;180)	(25.7;-163.66)
j_6				(21.4;53.1)
j_7				(40.8;-117)
j_8				(3.3;-78.8)
j_9				(14.9;180)

Table C.3: Sources weights j_p for the application of Chapter 8. Values in the table are $(|j_p|\, A_\delta$ [mA]; $\angle j_p$ [deg]$)$