



HAL
open science

Computational strategies for impedance boundary condition integral equations in frequency and time domains

Alexandre Dély

► **To cite this version:**

Alexandre Dély. Computational strategies for impedance boundary condition integral equations in frequency and time domains. Computational Physics [physics.comp-ph]. Ecole nationale supérieure Mines-Télécom Atlantique, 2019. English. NNT : 2019IMTA0135 . tel-02509607

HAL Id: tel-02509607

<https://theses.hal.science/tel-02509607>

Submitted on 17 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPERIEURE MINES-TELECOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE
COMUE UNIVERSITE BRETAGNE LOIRE
ECOLE DOCTORALE N°601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Génie électrique*
En cotutelle avec l'Université de Nottingham

Par **Alexandre DÉLY**

Computational Strategies for Impedance Boundary Condition Integral Equations in Frequency and Time Domains

Thèse présentée et soutenue à Paris le 15 mars 2019
Unité de recherche : Lab-STICC MOM-PIM
Thèse N° : 2019IMTA0135

Rapporteurs avant soutenance :

Trevor BENSON
Giuseppe VECCHI

Professor, University of Nottingham
Professore ordinario, Politecnico di Torino

Composition du Jury :

Président : Philippe POULIGUEN
Examineurs : Trevor BENSON
Francesco FERRANTI
Samuel NOSAL
Giuseppe VECCHI
Dir. de thèse : Francesco ANDRIULLI
Co-dir. de thèse : Kristof COOLS

Directeur de la recherche scientifique, Direction Générale de l'Armement
Professor, University of Nottingham
Maître de conférences, IMT Atlantique
Ingénieur-chercheur, Thales
Professore ordinario, Politecnico di Torino
Professore ordinario, Politecnico di Torino
Assistant professor, Delft University of Technology

PhD Thesis

Computational Strategies for Impedance Boundary Condition Integral Equations in Frequency and Time Domains

By **Alexandre DÉLY**

IMT Atlantique, Brest
Département Micro-ondes
Lab-STICC, MOM

University of Nottingham, Nottingham
Department of Electrical and Electronic Engineering
GGIEMR

Defended on 15/03/2019

Supervised by Francesco P. ANDRIULLI and Kristof COOLS

Jury

Francesco P. ANDRIULLI, Professore ordinario, Politecnico di Torino

Trevor BENSON, Professor, University of Nottingham

Kristof COOLS, Assistant professor, Delft University of Technology

Francesco FERRANTI, Maître de conférences, IMT Atlantique

Samuel NOSAL, Ingénieur-chercheur, Thales

Philippe POULIGUEN, Directeur de la recherche scientifique, Direction Générale de l'Armement

Giuseppe VECCHI, Professore ordinario, Politecnico di Torino

École nationale supérieure Mines-Télécom Atlantique Bretagne-Pays de la Loire
École doctorale MathSTIC

Submitted to the University of Nottingham for the degree of Doctor of Philosophy
Faculty of Engineering

Abstract

The Electric Field Integral Equation (EFIE) is widely used to solve wave scattering problems in electromagnetics using the Boundary Element Method (BEM). In the frequency domain, the linear systems stemming from the BEM suffer, amongst others, from two ill-conditioning problems: the low frequency breakdown and the dense mesh breakdown. Consequently, the iterative solvers require more iterations to converge to the solution, or they do not converge at all in the worst cases. These breakdowns are also present in the time domain, in addition to the DC instability which causes the solution to be completely wrong in the late time steps of the simulations. The time discretization is achieved using a convolution quadrature based on Implicit Runge-Kutta (IRK) methods, which yields a system that is solved by Marching-On-in-Time (MOT).

In this thesis, several integral equations formulations, involving Impedance Boundary Conditions (IBC) for most of them, are derived and subsequently preconditioned. In a first part dedicated to the frequency domain, the IBC-EFIE is stabilized for the low frequency and dense meshes by leveraging the quasi-Helmholtz projectors and a Calderón-like preconditioning. Then, a new IBC is introduced to enable the development of a multiplicative preconditioner for the new IBC-EFIE. In the second part on the time domain, the EFIE is regularized for the Perfect Electric Conductor (PEC) case, to make it stable in the large time step regime and immune to the DC instability. Finally, the solution of the time domain IBC-EFIE is investigated by developing an efficient solution scheme and by stabilizing the equation for large time steps and dense meshes.

Résumé

L'équation intégrale du champ électrique (EFIE) est très utilisée pour résoudre des problèmes de diffusion d'ondes électromagnétiques grâce à la méthode aux éléments de frontière (BEM). En domaine fréquentiel, les systèmes matriciels émergeant de la BEM souffrent, entre autres, de deux problèmes de mauvais conditionnement : l'augmentation du nombre d'inconnues et la diminution de la fréquence entraînent l'accroissement du nombre de conditionnement. En conséquence, les solveurs itératifs requièrent plus d'itérations pour converger vers la solution, voire ne convergent pas du tout. En domaine temporel, ces problèmes sont également présents, en plus de l'instabilité DC qui entraîne une solution erronée en fin de simulation. La discrétisation en temps est obtenue grâce à une quadrature de convolution basée sur les méthodes de Runge-Kutta implicites.

Dans cette thèse, diverses formulations d'équations intégrales utilisant notamment des conditions d'impédance aux frontières (IBC) sont étudiées et préconditionnées. Dans une première partie en domaine fréquentiel, l'IBC-EFIE est stabilisée pour les basses fréquences et les maillages denses grâce aux projecteurs quasi-Helmholtz et à un préconditionnement de type Calderón. Puis une nouvelle forme d'IBC est introduite, ce qui permet la construction d'un préconditionneur multiplicatif. Dans la seconde partie en domaine temporel, l'EFIE est d'abord régularisée pour le cas d'un conducteur électrique parfait (PEC), la rendant stable pour les pas de temps larges et immunisée à l'instabilité DC. Enfin, une résolution efficace de l'IBC-EFIE est recherchée, avant de stabiliser l'équation pour les pas de temps larges et les maillages denses.

Contents

Abstract	ii
Résumé	iii
Contents	viii
Preface	ix
Résumé long	xi
Acknowledgments	xv
I Frequency domain	1
1 Background and notations	2
1.1 Introduction to computational electromagnetics	2
1.2 Surface integral equations	4
1.2.1 Maxwell's equations	4
1.2.2 Time harmonic Maxwell's equations	5
1.2.3 Potentials in free space	6
1.2.4 Potentials for the magnetic sources	8
1.2.5 Far field	10
1.2.5.1 Far field approximation	10
1.2.5.2 Expression of the far fields	10
1.2.5.3 Radar cross section	12
1.2.6 Boundary conditions on discontinuous interfaces	12
1.2.6.1 Integral forms of Maxwell's equations	12
1.2.6.2 Boundary condition on the electric charges	13
1.2.6.3 Boundary condition on the electric currents	14
1.2.6.4 Boundary conditions for the magnetic quantities	15
1.2.7 Surface integral equations	15
1.2.7.1 Surface equivalence principle	15
1.2.7.2 Surface integral equations	17
1.2.7.3 Surface integral operators	19
1.2.7.4 Surface integral equations for the interior domain	20

1.3	Boundary element method	20
1.3.1	Method of moments	20
1.3.2	Discretization of the composition of operators	22
1.3.3	Basis functions	23
1.3.3.1	Rao-Wilton-Glisson basis functions	23
1.3.3.2	Buffa-Christiansen basis functions	25
1.3.3.3	Star basis functions	26
1.3.3.4	Loop basis functions	27
1.3.3.5	Global loops	28
1.3.4	Discretization of the surface integral operators	29
1.4	Impedance boundary conditions	30
1.5	Preconditioning	33
1.5.1	Condition number and preconditioning	34
1.5.1.1	Condition number	34
1.5.1.2	Preconditioning and iterative solvers	35
1.5.2	Analysis of the operators on a sphere	37
1.5.2.1	Spectrum of operators	37
1.5.2.2	Low frequency breakdown	40
1.5.2.3	Dense mesh breakdown	42
1.5.2.4	Spurious resonances (high frequency)	44
1.5.2.5	Numerical cancellations (low frequency)	46
1.5.3	Solutions to the ill-conditioning	48
1.5.3.1	Calderón preconditioning	48
1.5.3.2	Loop-star decomposition	50
1.5.3.3	Quasi-Helmholtz projectors	55
2	Low frequency and dense mesh stable IBC-EFIE	62
2.1	Introduction	62
2.2	Impedance Boundary Condition - Electric Field Integral Equation . .	63
2.2.1	Integral equation of the IBC-EFIE	63
2.2.2	Discretization of the IBC-EFIE	64
2.3	Analysis of the IBC-EFIE operator	65
2.3.1	Spectrum of the IBC-EFIE operator	65
2.3.2	Analytic solution of the IBC-EFIE on a sphere	66
2.3.3	Low frequency behaviour	68
2.3.3.1	Low frequency breakdown	68
2.3.3.2	Numerical cancellations	69
2.3.4	Analysis of the dense mesh breakdown	69
2.4	Solution of the low frequency problems	70
2.5	Solution of the dense mesh breakdown	73
2.6	Implementation details	73
2.7	Numerical results	74
2.7.1	Correctness of the formulation	75
2.7.2	Stability of the formulation	77
2.7.3	Realistic case scenario	79

2.8	Conclusion and future work	80
3	Multiplicative preconditioner for a new IBC-EFIE	83
3.1	Introduction	83
3.2	Background and notations	84
3.2.1	Construction of the IBC-EFIE	84
3.2.2	Discretization of the IBC-EFIE	85
3.3	New Impedance Boundary Condition	86
3.3.1	Definition of the new IBC	86
3.3.2	Validity of the new IBC	87
3.3.3	Accuracy of the new IBC	88
3.4	Vector Spherical Harmonics analysis	88
3.4.1	Vector Spherical Harmonics expansion of the new IBC operator	89
3.4.2	Analytic solution for the currents	90
3.4.3	Low frequency behaviour	91
3.4.4	Dense mesh behaviour	92
3.5	Construction of a Calderón multiplicative preconditioner	93
3.6	Numerical results	94
3.6.1	Results related to the new IBC accuracy	94
3.6.2	Results related to the stability of the formulation	96
3.6.3	Realistic application	100
3.7	Conclusion and future work	100
II	Time domain	104
4	Introduction to the time domain	105
4.1	Surface integral equations in the time domain	105
4.2	Time domain discretization and marching-on-in-time	108
4.3	Implicit Runge-Kutta convolution quadrature	112
4.3.1	Laplace transform and Z-transform	112
4.3.1.1	Laplace transform	112
4.3.1.2	Z-transform	113
4.3.2	Runge-Kutta methods	114
4.3.3	Convolution quadrature using IRK methods	116
4.3.4	Discretization of time domain integral equations	117
4.3.5	Computation of the matrix elements	120
4.3.5.1	Operators in the Laplace domain	120
4.3.5.2	Computation of the matrix elements in the Z-domain	121
4.3.5.3	Kronecker product	121
4.4	DC instability	122

5	Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods	126
5.1	Introduction	126
5.2	Background and notations	128
5.2.1	Perfect Electric Conductor - Electric Field Integral Equation	128
5.2.1.1	PEC-EFIE in the time domain	128
5.2.1.2	PEC-EFIE in the Laplace domain	128
5.2.2	Space Galerkin discretization	129
5.2.3	Time discretization with IRK convolution quadrature	129
5.3	Analysis of the large time step breakdown and the DC instability	131
5.3.1	Large time step breakdown	131
5.3.2	DC instability	132
5.3.2.1	DC instability in the PEC-EFIE	132
5.3.2.2	DC instability in the differentiated PEC-EFIE	133
5.4	Regularized PEC-EFIE	134
5.4.1	Regularization in the Laplace domain	134
5.4.2	Time discretization of the regularized system	135
5.4.2.1	Interaction matrices	135
5.4.2.2	Excitation vector (analytic evaluation)	136
5.4.2.3	Excitation vector (numerical evaluation)	136
5.4.2.4	Solution	137
5.5	Implementation details	138
5.5.1	Quasi-Helmholtz projectors	138
5.5.2	Excitation vectors	138
5.5.3	Retrieving the original current from the auxiliary current	139
5.5.4	Marching-on-in-time	139
5.5.5	Interaction matrix multiplication by a vector	139
5.6	Numerical results	140
5.6.1	Results related to the DC instability	141
5.6.2	Results related to the large time step breakdown	143
5.7	Conclusion and future work	146
6	Large time step and dense mesh stable time domain IBC-EFIE	147
6.1	Introduction	147
6.2	Background and notations	148
6.2.1	Integral equations in the time and Laplace domains	148
6.2.2	Spacial discretization	149
6.2.3	Time discretization	150
6.2.3.1	Discretization of the vectors	150
6.2.3.2	Discretization of the operators	151
6.2.3.3	Discretization of the equations	152
6.3	Solution of the time differentiated IBC-EFIE	153
6.3.1	Overview of the problems	153
6.3.1.1	Solving for the two currents simultaneously	154
6.3.1.2	Substitution of the magnetic current	154

6.3.2	Marching-on-in-time	155
6.3.3	Computation of the auxiliary current	156
6.4	Preconditioning of the time domain IBC-EFIE	157
6.4.1	Overview of the problems	157
6.4.2	Large time step stabilization	158
6.4.2.1	Left hand side	158
6.4.2.2	Right hand side	160
6.4.3	Dense mesh stabilization	162
6.5	Numerical results	163
6.5.1	Correctness of the formulation	163
6.5.2	Stability of the formulation	164
6.6	Conclusion and future work	165
	Conclusions and future work	169
	Appendix	170
A	Computation of the MoM matrix elements	171
A.1	Expression of the matrix elements	171
A.2	Integration rules	172
A.3	Singularity extraction	174
	Acronyms and notations	177
	Publications	179
	Bibliography	180

Preface

The Electric Field Integral Equation (EFIE) is commonly used to solve scattering problems in electromagnetics using the Boundary Element Method (BEM). Unfortunately, in the frequency domain (time harmonics), the linear system emanating from the BEM discretization suffers from the dense mesh and the low frequency breakdowns. In other words, when the frequency is decreased or when the number of unknowns is increased, the condition number of the system increases accordingly. This ill-conditioning increases the number of iterations required by iterative solvers to converge to the solution, and in the most extreme cases it prevents them from converging.

The first part of this thesis focuses on frequency domain integral equations. The first introductory chapter presents the state-of-the-art of computational electromagnetics that is necessary to understand the work done. In particular, the surface integral equations are derived from the Maxwell's equations, the BEM as well as the basis functions are detailed. Also, the ill-conditioning problems that are encountered and the preconditioning techniques that are used in this work are reviewed.

In the second chapter, the Leontovich Impedance Boundary Condition (IBC) is used in conjunction to the EFIE. It results in the IBC-EFIE which suffers from the breakdowns described above. Using recent preconditioning techniques, a stable IBC-EFIE is constructed: the new formulation is well-conditioned and yields the correct solution for arbitrary low frequency and arbitrary dense meshes, both on simply connected and multiply connected geometries. In particular, to solve the low frequency breakdown, the quasi-Helmholtz projectors are used to separate and rescale independently the two Helmholtz components of the fields in the equation. Then, the unbounded part of the spectrum of the equation, that is responsible of the dense mesh breakdown, is regularized by leveraging a Calderón-like preconditioning. To obtain a conforming scheme, the magnetic currents sources are discretized using dual basis functions.

Then in the third chapter, a new IBC is introduced to build an IBC-EFIE that can be preconditioned in a multiplicative manner. In the limit of a Perfect Electric Conductor (PEC), the new formulation tends to the classical Calderón multiplicative preconditioner. To contrast with the first one developed in the precedent chapter, this second IBC-EFIE formulation does not use quasi-Helmholtz projectors which makes it relatively easy to implement.

In the second part of this thesis, time domain integral equations are investigated. The fourth chapter is an introduction that is specific to the time domain. The spacial

and temporal discretizations yield a system that can be solved with a Marching-On-in-Time (MOT) algorithm. In addition to the dense mesh breakdown and the large time step breakdowns (time domain analogue of the low frequency breakdown), the formulations are plagued by the DC instability. It corresponds to the existence of a non physical static current in the solution of the equation which produces a wrong result in the late time steps. Instead of the classical testing with time basis functions, the temporal discretization can be achieved using a convolution quadrature based on Implicit Runge-Kutta (IRK) methods.

In the fifth chapter, in the context of a time discretization with an IRK convolution quadrature, it is explained how the quasi-Helmholtz projectors can be used to regularize the PEC-EFIE. The new formulation is made stable for large time steps and immune to the DC instability, while maintaining the compatibility with existing fast solvers.

Finally in the sixth chapter, the solution of the IBC-EFIE is investigated in the time domain. Even though a naive discretization of the equation results in a suboptimal quadratic complexity of the MOT scheme, an efficient MOT is obtained by taking advantage of the particular form of the IBC-EFIE operator. Then, using the quasi-Helmholtz projectors those effectiveness has been established on the PEC-EFIE in the precedent chapter, a stabilization strategy is developed for the new MOT scheme by making use of the preconditioners built for the frequency domain IBC-EFIE. It results in a large time step and dense mesh stable solver for the time domain IBC-EFIE that is computationally efficient.

Résumé long

L'équation intégrale du champ électrique (EFIE) est couramment utilisée pour résoudre des problèmes de diffusion d'ondes électromagnétiques grâce à la méthode aux éléments de frontière (BEM). Cette méthode consiste à reformuler le problème de diffusion dans tout l'espace en terme d'une équation intégrale sur la surface de l'objet diffusant les ondes. Puis, le problème est discretisé en approximant la surface par un maillage triangulaire sur lequel sont définies les fonctions de bases Rao-Wilton-Glisson (RWG) et Buffa-Christiansen (BC). En domaine fréquentiel, le système d'équations linéaires qui émerge de la discrétisation de la BEM souffre, entre autres, de deux problèmes de mauvais conditionnement. D'une part lorsque la fréquence de simulation diminue, et d'autre part lorsque le nombre d'inconnues augmente, le nombre de conditionnement du système augmente en conséquence. Ainsi, les solveurs itératifs ont besoin de plus d'itérations pour converger vers la solution ce qui rallonge le temps de résolution, ou ils peuvent ne pas converger du tout dans des cas plus extrêmes. Le but de cette thèse est d'étudier ces problèmes de mauvais conditionnement dans plusieurs cas dans les domaines fréquentiel et temporel afin de les résoudre grâce à différentes stratégies de préconditionnement. La première partie traite du domaine fréquentiel et la seconde partie le domaine temporel. Chaque partie comporte trois chapitres dont une introduction et les développements de deux formulations.

Le premier chapitre de cette thèse est introductif. Les équations intégrales sont d'abord dérivées à partir des équations de Maxwell. Ensuite, les fonctions de bases mises en jeu sont présentées. Enfin, les problèmes de préconditionnement ainsi que certaines de leurs solutions sont expliqués dans le cas d'un conducteur électrique idéal (PEC).

Les conditions d'impédance aux limites (IBC) sont très utilisées pour modéliser les matériaux impénétrables avec pertes. En notant \mathbf{J}_s la densité de courant électrique de surface, \mathbf{M}_s la densité de courant magnétique de surface, z^{imp} l'impédance et $\hat{\mathbf{n}}$ la normale à la surface, l'IBC classique de Leontovich se note

$$\mathbf{M}_s = -z^{\text{imp}} \hat{\mathbf{n}} \times \mathbf{J}_s. \quad (1)$$

Lorsque les IBC sont combinées à l'EFIE, il en résulte l'IBC-EFIE qui peut s'écrire de façon pratique

$$\mathcal{S} \mathbf{J}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \quad (2)$$

où \mathbf{E}^{inc} est le champ électrique incident et \mathcal{S} est l'opérateur de l'IBC-EFIE. Afin de discrétiser conformément l'équation, les densités de courants électriques de surface sont représentés par les fonctions de base RWG, tandis que les courants magnétiques

sont représentés par des fonctions de base BC. Une fois l'équation discrétisée, le système linéaire à résoudre est

$$\mathbf{S}\mathbf{J} = \mathbf{E}. \quad (3)$$

Ce système souffre des problèmes de conditionnement décrits plus haut. Dans le deuxième chapitre, ce problème est étudié analytiquement sur une sphère en utilisant les harmoniques sphériques vectorielles (VSH). Il est montré que contrairement au conducteur électrique idéal pour lequel la PEC-EFIE a un nombre de conditionnement qui croît quadratiquement avec l'inverse de la fréquence et l'inverse de la taille moyenne des arêtes du maillage, dans le cas de l'IBC-EFIE l'accroissement est seulement linéaire en général mais il dépend également de l'impédance. Ensuite une IBC-EFIE stable est développée. Cette nouvelle formulation est bien conditionnée et délivre la bonne solution pour des fréquences arbitrairement basses et des maillages arbitrairement denses, à la fois sur des géométries à connexité simple ou multiple. Plus précisément, l'utilisation des projecteurs quasi-Helmholtz pour séparer et redimensionner indépendamment les deux composantes de la décomposition de Helmholtz des champs, permet d'obtenir une formulation stable à basse fréquence. Celle-ci s'écrit

$$\mathbf{L}\mathbf{S}\mathbf{R}\mathbf{Y} = \mathbf{L}\mathbf{E}, \quad (4)$$

où \mathbf{L} est le préconditionneur à gauche qui a pour but d'empêcher l'annulation de la partie solénoïdale de l'excitation, \mathbf{R} est le préconditionneur à droite qui a pour but d'empêcher les annulations numériques dans le courant et \mathbf{Y} est une inconnue auxiliaire pour laquelle le système préconditionné est résolu ($\mathbf{J} = \mathbf{R}\mathbf{Y}$). Puis, un préconditionnement de type Calderón utilisant les propriétés spectrales des opérateurs permet de régulariser la partie non bornée du spectre. Ainsi, la formulation stable à basse fréquence et pour les maillages denses est

$$\tilde{\mathbf{T}}\mathbf{G}_m^{-1}\mathbf{L}\mathbf{S}\mathbf{R}\mathbf{Y} = \tilde{\mathbf{T}}\mathbf{G}_m^{-1}\mathbf{L}\mathbf{E}, \quad (5)$$

où $\tilde{\mathbf{T}}$ est le préconditionneur discrétisé avec les fonctions BC et \mathbf{G}_m^{-1} est l'inverse de la matrice de Gram mixte permettant le passage entre les discrétisations avec des fonctions RWG et BC.

Dans le troisième chapitre, une nouvelle forme d'IBC est introduite dans le but de construire une IBC-EFIE qui peut être préconditionnée multiplicativement. La nouvelle IBC se note

$$\mathbf{M}_s = -z^{\text{imp}}\mathcal{N}_\delta\mathbf{J}_s \quad (6)$$

où \mathcal{N}_δ est un opérateur défini à partir de la partie singulière \mathcal{T}_s de l'opérateur de la PEC-EFIE avec un nombre d'onde complexe:

$$\mathcal{N}_\delta = 2ik\mathcal{T}_s|_{ik=\frac{1+i}{\delta}}. \quad (7)$$

L'idée est de remplacer l'IBC de Leontovich par une IBC différente afin que, d'une part, la solution du problème reste inchangée dans une marge de tolérance contrôlable grâce au paramètre δ s'apparentant à la distance caractéristique de l'effet de peau, et d'autre part, la nouvelle IBC modifie qualitativement les propriétés spectrales de

l'opérateur de l'IBC-EFIE, ce qui le rend plus simple et plus naturel à préconditionner. Une étude du nouvel opérateur basées sur les VSH permet d'obtenir la solution analytique sur une sphère d'une part, et le comportement de la nouvelle IBC-EFIE à basse fréquence et pour les maillages denses d'autre part. Ainsi le système préconditionné s'écrit

$$\tilde{\mathbf{T}}\mathbf{G}_m^{-1}\mathbf{S}\mathbf{J} = \tilde{\mathbf{T}}\mathbf{G}_m^{-1}\mathbf{E} \quad (8)$$

où le préconditionneur discrétisé avec les fonctions BC est

$$\tilde{\mathbf{T}} = -ik\mathbb{T}_s + \frac{1}{ik + \frac{1+i}{\delta} \frac{z^{\text{imp}}}{\eta}} \mathbb{T}_h. \quad (9)$$

Dans la limite du conducteur électrique idéal (PEC) i.e. lorsque $z^{\text{imp}} \rightarrow 0$, la nouvelle formulation se réduit à la PEC-EFIE munie de son preconditionneur multiplicatif classique. Cette seconde formulation se différencie aussi de la première par la non-utilisation des projecteurs quasi-Helmholtz, ce qui la rend plus simple à implémenter.

La seconde partie de cette thèse s'intéresse à la résolution d'équations intégrales en domaine temporel. Les discrétisations en espace et en temps produisent un système qui peut être résolu par une procédure de marche dans le temps (MOT). En plus des problèmes de maillages denses et de pas de temps larges (l'analogie des basses fréquences en domaine temporel), les formulations souffrent aussi d'instabilités DC. Elles correspondent à l'existence d'un courant statique non-physique dans la solution de l'équation ce qui donne un résultat complètement erroné dans les derniers temps de la simulation. Au lieu de tester l'équation avec les fonctions de base temporelles comme il est fait classiquement, la discrétisation en temps peut aussi être obtenue par le biais d'une quadrature de convolution basée sur les méthodes de Runge-Kutta implicites. Un avantage pratique de cette technique est qu'il est relativement simple d'obtenir un solveur en domaine temporel à partir d'un solveur en domaine fréquentiel. Tout ceci est présenté dans le quatrième chapitre qui est un chapitre introductif dédié uniquement au domaine temporel.

Dans le cadre d'une discrétisation en temps basée sur les méthodes de Runge-Kutta implicites, le cinquième chapitre explique comment les projecteurs quasi-Helmholtz peuvent être utilisés pour régulariser la PEC-EFIE. La stabilisation est en fait réalisée dans le domaine de Laplace où la matrice d'interactions $\mathbf{Z}(s)$ est préconditionnée à gauche par $\mathbf{L}(s)$ et à droite par $\mathbf{R}(s)$. La matrice préconditionnée est notée

$$\mathbf{Z}^{\text{reg}}(s) = \mathbf{L}(s)\mathbf{Z}(s)\mathbf{R}(s). \quad (10)$$

En suivant la procédure de discrétisation temporelle utilisant la quadrature de convolution basée sur les méthodes de Runge-Kutta, la MOT obtenue est donnée par la formule

$$\mathbf{z}_0^{\text{reg}}\mathbf{Y}_i = \mathbf{V}_i - \sum_{j=1}^{N_{\text{conv}}} \mathbf{z}_j^{\text{reg}}\mathbf{Y}_{i-j} \quad (11)$$

où $(\mathbf{z}_j^{\text{reg}})_j$ est la séquence de matrices d'interactions dont le premier élément $\mathbf{z}_0^{\text{reg}}$ a un nombre de conditionnement stable lorsque le pas de temps $\Delta t \rightarrow +\infty$, $(\mathbf{Y}_i)_i$ est

une séquence de vecteurs obtenue par la MOT étant utilisée pour récupérer la densité de courant du problème original, et $(\mathbf{V}_i)_i$ est la séquence de vecteurs d'excitation conforme à la stabilisation effectuée sur la formulation. Ainsi, l'équation est rendue stable pour des pas de temps larges et immunisée à l'instabilité DC. Ceci est fait tout en maintenant la compatibilité avec les solveurs rapides existants.

Enfin, la résolution de l'IBC-EFIE en domaine temporel est étudiée dans le sixième chapitre. Une discrétisation naïve de l'équation résulte en une procédure de MOT sous-optimale dont la complexité algorithmique est quadratique. Ceci est dû aux pertes introduites par l'impédance qui se retrouvent dans les matrices d'interactions et par conséquent, leur nombre n'est pas borné. Il est montré qu'une MOT efficace peut être obtenue en tirant avantage de la forme particulière de l'opérateur de l'IBC-EFIE. En effet, au lieu d'avoir l'impédance dans les matrices d'interactions, il est possible de la déplacer dans une séquence de vecteurs auxiliaires

$$\mathbf{M}'_i = -\tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \sum_{j=1}^i \tilde{\mathbf{z}}_j^{\text{imp}} \mathbf{J}_{i-j} \quad (12)$$

qui sont calculés rapidement à chaque pas de temps grâce à un algorithme basé sur la transformée de Fourier rapide (FFT). Ceci permet alors d'obtenir une MOT efficace n'utilisant que les matrices d'interactions du conducteur électrique idéal:

$$\mathbf{S}_0^{\text{der}} \mathbf{J}_i = \mathbf{E}_i^{\text{der}} - \eta \sum_{j=1}^i \mathbf{T}_j^{\text{der}} \mathbf{J}_{i-j} - \sum_{j=1}^i \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \right) \mathbf{M}_{i-j} - \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_0 - \mathbf{K}_0^{\text{der}} \right) \mathbf{M}'_i. \quad (13)$$

Puis, en utilisant les projecteurs quasi-Helmholtz dont l'efficacité a été prouvée sur la PEC-EFIE dans le chapitre précédent, une stratégie de stabilisation est développée pour la nouvelle MOT en réutilisant les préconditionneurs construits pour l'IBC-EFIE en domaine fréquentiel. Il en résulte un solveur pour l'IBC-EFIE en domaine temporel ayant une complexité asymptotique efficace, et étant stable pour les pas de temps larges et les maillages denses.

Acknowledgments

I would like to thank my supervisors Francesco Andriulli and Kristof Cools for giving me this wonderful opportunity of doing research. This accomplishment would have not been possible without their constant support and their immense knowledge.

During these three years, I had the occasion to visit and work in different places. In particular, I would like to thank the members of Microwave Department of IMT Atlantique in Brest, in addition to the members of the GGIEMR group in the University of Nottingham Department of Electrical and Electronic Engineering.

I also express my gratitude to the DGA and the DSTL that funded my thesis. The France-United Kingdom research collaboration framework in which this thesis fits has been a truly rewarding experience.

I also thank the other PhD students and postdocs of the lab, Adrien, Axelle, Clément, Davide, François, John Erick, Lyes, Maxime, Rajendra, Simon and Tiffany.

Last but not least, I also thank my family, especially my parents and my sisters who have always supported me.

To my parents.

Part I
Frequency domain

Chapter 1

Background and notations

This chapter introduces the basic concepts of computational electromagnetics that are necessary to comprehend the work presented in this thesis. First, the surface integral equations to be solved are derived from the Maxwell's equations in the frequency domain. Then, the Boundary Element Method (BEM) and the basis functions used for the spacial discretization of the integral equations are detailed. A section is dedicated to the presentation of the Impedance Boundary Conditions (IBC) that have a prevalent presence in this work. The last section is an overview of the problems of ill-conditioning that can arise in the linear systems originating from the Boundary Element Method (BEM), and a presentation of the techniques of preconditioning used to treat these problems.

1.1 Introduction to computational electromagnetics

Computational electromagnetics is the science of solving the Maxwell's equations numerically. It is a vast topic of active research that aims to improve the existing solving algorithms and formulations on one or several features including:

- The computational complexity in time and memory. The Boundary Element Method (BEM) that is used in this work yields a linear system of N equations and N unknowns that would have a quadratic ($\mathcal{O}(N^2)$) complexity in space (memory) and a cubic ($\mathcal{O}(N^3)$) complexity in execution time, if it was solved naively. Since computers have a limited processing power and a limited memory, an algorithm is usually considered to be fast or scalable if its asymptotic complexity is linear, up to a logarithmic factor ($\mathcal{O}(N\log(N)^k)$). In this work, formulations compatible with this latter range of complexity are looked for.
- The parallelization. With the increasing number of cores in processors and the increasing number of nodes in supercomputers, the scalability of the parallelization is a topic that is becoming more and more ubiquitous.

- The precision. Computers work with floating point numbers that have a finite precision. Therefore, the solvers need to ensure that a small error of approximation at the beginning of an algorithm does not yield a large error in the result at the end.
- The accuracy. Somehow linked to the precision, it refers to the accurate computation of a quantity with the maximum or a given number of correct digits.
- The stability. It is related to the behaviour of a scheme when a given parameter (frequency, number of time steps, mesh density, physical parameters such as the impedance, etc.) is changed. Stable formulations tend to have the same computational cost and precision, independently of the parameters of the simulation.

These features are linked to each other and usually trade-offs have to be done to privilege one over the others. The most obvious trade-off is with the computational cost, but estimating the accuracy, the precision or the stability of an algorithm is also an important topic of research that requires advanced mathematical tools in numerical error analysis.

The focus of this work is on preconditioning. The starting point of preconditioning is an original formulation that gives rise to an ill-conditioned linear system. For example, the condition number of the system increases when the mesh corresponding to the geometry of the scatterer is refined (dense mesh breakdown) or when the frequency decreases (low frequency breakdown). As iterative solvers are used to solve the linear system, in general the larger is the condition number, the more iterations are needed to converge to the solution. The goal of preconditioning is to multiply the original matrix by another one such that the condition number of the product is lower than the condition number of the original matrix, and as consequence it reduces the number of iterations required to solve the problem.

By involving additional quantities to be computed compared to the original formulations, the preconditioned formulations seem more complicated. But overall, the preconditioning improves the time complexity by reducing the number of iterations required for the convergence to the solution. It also improves the precision because a low condition number means that a small error in the input will not result in a large error in the solution. In this work, the preconditioning aims to stabilize the different formulations such that the condition numbers of the systems remain bounded, for example in the low frequency and dense mesh limits.

1.2 Surface integral equations

1.2.1 Maxwell's equations

The starting point of the electromagnetic theory is the set of Maxwell's equations. The local equations at a position \mathbf{r} and time t read

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \rho(\mathbf{r}, t) \quad \text{Maxwell-Gauss} \quad (1.1a)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad \text{Maxwell-Thomson} \quad (1.1b)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad \text{Maxwell-Faraday} \quad (1.1c)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} + \mathbf{J}(\mathbf{r}, t) \quad \text{Maxwell-Ampère} \quad (1.1d)$$

where

- \mathbf{E} (V/m) is the electric field intensity,
- \mathbf{H} (A/m) is the magnetic field intensity,
- \mathbf{D} (C/m²) is the electric displacement field,
- \mathbf{B} (Wb/m² or T) is the magnetic flux density,
- ρ (C/m³) is the electric charge density,
- \mathbf{J} (A/m²) is the electric current density.

The Maxwell's equations are completed with the constitutive equations

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (1.2a)$$

$$\mathbf{B} = \mu \mathbf{H} \quad (1.2b)$$

where

- ε (F/m) is the permittivity of the medium,
- μ (H/m) is the permeability of the medium.

A priori, ε and μ depend on the position (inhomogeneous medium), on the fields (non-linear medium), on time (non-stationary medium) or frequency (dispersive medium), and may be tensors (anisotropic medium). However, in this work it is assumed that the medium has none of these complicating properties so that the permittivity and the permeability are constant scalars, which is a common practical assumption for wave propagation in a wide range of media including air and vacuum.

The permeability and permittivity of vacuum are defined exactly to be

$$\mu_0 = 4\pi 10^{-7} \text{ H/m} \approx 1.26 \times 10^{-6} \text{ H/m} \quad (1.3a)$$

$$\varepsilon_0 = \frac{1}{\mu_0 c_0^2} \approx 8.85 \times 10^{-12} \text{ F/m} \quad (1.3b)$$

where c_0 is the speed of light in vacuum, also defined exactly as

$$c_0 = 299792458 \text{ m/s.} \quad (1.4)$$

Finally, the characteristic impedance of vacuum is exactly

$$\eta_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} = \mu_0 c_0 = 119.9169832\pi \ \Omega \approx 376.73 \ \Omega. \quad (1.5)$$

1.2.2 Time harmonic Maxwell's equations

Define $\mathcal{F}(X)$ the Fourier transform of an \mathbb{R} integrable function X that may be scalar or vector. Here f is the frequency and $\omega = 2\pi f$ is the angular frequency.

$$\mathcal{F}(X)(\omega) = \int_{t \in \mathbb{R}} X(t) e^{-i\omega t} dt. \quad (1.6)$$

To simplify the notations the same symbol is used for a function and its Fourier transform, the distinction being clear from their arguments, so that $\mathcal{F}(X)(\omega)$ is noted $X(\omega)$ when there is no possible confusion. Using the property of the Fourier transform

$$\mathcal{F}\left(\frac{\partial X}{\partial t}\right)(\omega) = i\omega \mathcal{F}(X)(\omega) \quad (1.7)$$

the Maxwell's equations are rewritten in the frequency domain (Fourier domain)

$$\nabla \cdot \mathbf{D}(\mathbf{r}, \omega) = \rho(\mathbf{r}, \omega) \quad \text{Maxwell-Gauss} \quad (1.8a)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, \omega) = 0 \quad \text{Maxwell-Thomson} \quad (1.8b)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, \omega) = -i\omega \mathbf{B}(\mathbf{r}, \omega) \quad \text{Maxwell-Faraday} \quad (1.8c)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, \omega) = i\omega \mathbf{D}(\mathbf{r}, \omega) + \mathbf{J}(\mathbf{r}, \omega) \quad \text{Maxwell-Ampère.} \quad (1.8d)$$

Here the time-dependent quantities are real valued and the frequency-dependent quantities are complex valued. It is very practical to express the Maxwell's equations and solve them in the frequency domain, because it usually simplifies many operations such as the time derivative, the time integration or the time delay.

Usually, a single frequency is used so that any time dependent quantity is sinusoidal. At an angular frequency ω , any sinusoidal signal can be written in the form

$$X(t) = |X(\omega)| \cos(\omega t + \phi(\omega)) \quad (1.9a)$$

$$= \Re(|X(\omega)| e^{i\phi(\omega)} e^{i\omega t}) \quad (1.9b)$$

$$= \Re(X(\omega) e^{i\omega t}) \quad (1.9c)$$

1. Background and notations

where $|X(\omega)|$ is the amplitude and $\phi(\omega)$ is the phase of the phasor $X(\omega) = |X(\omega)| e^{i\phi(\omega)}$. For example, this scheme is applied on the Maxwell-Faraday equation (1.1c)

$$\Re(\nabla \times \mathbf{E}(\mathbf{r}, \omega) e^{i\omega t}) = -\frac{\partial}{\partial t} (\Re(\mathbf{B}(\mathbf{r}, \omega) e^{i\omega t})) \quad (1.10a)$$

$$= \Re(-i\omega \mathbf{B}(\mathbf{r}, \omega) e^{i\omega t}). \quad (1.10b)$$

By knowing that only the real part matters, \Re can be ignored in the computations. Then, also the time harmonic factor $e^{i\omega t}$ can be simplified which results in

$$\nabla \times \mathbf{E}(\mathbf{r}, \omega) = -i\omega \mathbf{B}(\mathbf{r}, \omega) \quad (1.11)$$

which is exactly the equation (1.8c) obtained with the Fourier transform. The main advantage is that it simplifies many computations with the fields by avoiding the cumbersome trigonometric functions. However, care must be taken in general as quantities that are the product of two fields in the time domain (e.g. energy or power related quantities such as the Poynting vector) do not fit in this scheme (since $\Re(a)\Re(b) \neq \Re(ab)$ in general for $a, b \in \mathbb{C}$).

Note that in the literature, numerous authors use the other convention by assuming a $e^{-i\omega t}$ factor, in which case, most results are converted into the other convention by flipping the sign in front of the imaginary unit ($+i \leftrightarrow -i$). In this work, the $e^{i\omega t}$ convention is used exclusively as described in this section.

1.2.3 Potentials in free space

In this section, the scalar and vector potentials are derived in the frequency domain at a given angular frequency ω . Assume that the sources in the Maxwell's equations, i.e. the charge density $\rho(\mathbf{r})$ and the current density $\mathbf{J}(\mathbf{r})$, are known. The goal is to express the scattered fields in terms of potentials that depends on the sources.

The Maxwell-Thomson equation (1.8b) is $\nabla \cdot \mathbf{B}(\mathbf{r}) = 0$, so it exists a vector field \mathbf{A} such that

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}). \quad (1.12)$$

\mathbf{A} is the vector potential (in V.s/m). It is not unique: its curl $\nabla \times \mathbf{A}$ has been set but its divergence $\nabla \cdot \mathbf{A}$ can still be freely chosen, which is done later with the choice of gauge. Inserting (1.12) in the Maxwell-Faraday equation (1.8c) results in

$$\nabla \times (\mathbf{E}(\mathbf{r}) + i\omega \mathbf{A}(\mathbf{r})) = \mathbf{0}. \quad (1.13)$$

So it exist a scalar field ϕ such that

$$\mathbf{E}(\mathbf{r}) + i\omega \mathbf{A}(\mathbf{r}) = -\nabla \phi(\mathbf{r}). \quad (1.14)$$

ϕ is the scalar potential (in V). It is also non-unique, all the possible choices differing by a constant field.

Then, the equations verified by the potentials are looked for. The equation for \mathbf{A} , is obtained by expressing $\mu \nabla \times \mathbf{H}$ in two different ways. On one hand using (1.8d), (1.2a) and (1.14),

$$\mu \nabla \times \mathbf{H}(\mathbf{r}) = i\omega\mu \mathbf{D}(\mathbf{r}) + \mu \mathbf{J}(\mathbf{r}) \quad (1.15a)$$

$$= i\omega\mu\varepsilon \mathbf{E}(\mathbf{r}) + \mu \mathbf{J}(\mathbf{r}) \quad (1.15b)$$

$$= i\omega\mu\varepsilon (-i\omega \mathbf{A}(\mathbf{r}) - \nabla\phi(\mathbf{r})) + \mu \mathbf{J}(\mathbf{r}) \quad (1.15c)$$

$$= \omega^2\mu\varepsilon \mathbf{A}(\mathbf{r}) - i\omega\mu\varepsilon \nabla\phi(\mathbf{r}) + \mu \mathbf{J}(\mathbf{r}). \quad (1.15d)$$

On the other hand using (1.2b) and (1.12),

$$\mu \nabla \times \mathbf{H}(\mathbf{r}) = \nabla \times \mathbf{B}(\mathbf{r}) \quad (1.16a)$$

$$= \nabla \times (\nabla \times \mathbf{A}(\mathbf{r})) \quad (1.16b)$$

$$= \nabla (\nabla \cdot \mathbf{A}(\mathbf{r})) - \nabla^2 \mathbf{A}(\mathbf{r}). \quad (1.16c)$$

Equating (1.15d) to (1.16c) and rearranging the terms results in

$$\nabla^2 \mathbf{A}(\mathbf{r}) + k^2 \mathbf{A}(\mathbf{r}) = -\mu \mathbf{J}(\mathbf{r}) + \nabla (i\omega\mu\varepsilon\phi(\mathbf{r}) + \nabla \cdot \mathbf{A}(\mathbf{r})) \quad (1.17)$$

where the wave number k (in m^{-1}) is introduced as

$$k = \omega\sqrt{\mu\varepsilon}. \quad (1.18)$$

Then, for the equation of ϕ , the quantity $\frac{1}{\varepsilon} \nabla \cdot \mathbf{D}$ is expressed in two different ways. On one hand using (1.8a),

$$\frac{1}{\varepsilon} \nabla \cdot \mathbf{D}(\mathbf{r}) = \frac{\rho(\mathbf{r})}{\varepsilon}. \quad (1.19)$$

On the other hand using (1.2a) and (1.14),

$$\frac{1}{\varepsilon} \nabla \cdot \mathbf{D}(\mathbf{r}) = \nabla \cdot \mathbf{E}(\mathbf{r}) \quad (1.20a)$$

$$= \nabla \cdot (-i\omega \mathbf{A}(\mathbf{r}) - \nabla\phi(\mathbf{r})) \quad (1.20b)$$

$$= -i\omega \nabla \cdot \mathbf{A}(\mathbf{r}) - \nabla^2 \phi(\mathbf{r}). \quad (1.20c)$$

Equating (1.19) to (1.20c), rearranging the terms, and adding $k^2\phi(\mathbf{r}) = -i\omega (i\omega\mu\varepsilon\phi(\mathbf{r}))$ on both sides results in

$$\nabla^2 \phi(\mathbf{r}) + k^2 \phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\varepsilon} - i\omega (i\omega\mu\varepsilon\phi(\mathbf{r}) + \nabla \cdot \mathbf{A}(\mathbf{r})). \quad (1.21)$$

From (1.12), $\nabla \times \mathbf{A}$ has been set but the choice of $\nabla \cdot \mathbf{A}$ is remaining. Regarding (1.17) and (1.21), it is clear that the natural choice for $\nabla \cdot \mathbf{A}$ is the Lorenz gauge

$$\nabla \cdot \mathbf{A}(\mathbf{r}) = -i\omega\mu\varepsilon\phi(\mathbf{r}). \quad (1.22)$$

1. Background and notations

With this gauge choice, (1.17) and (1.21) become

$$\nabla^2 \mathbf{A}(\mathbf{r}) + k^2 \mathbf{A}(\mathbf{r}) = -\mu \mathbf{J}(\mathbf{r}) \quad (1.23a)$$

$$\nabla^2 \phi(\mathbf{r}) + k^2 \phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\varepsilon}. \quad (1.23b)$$

The potentials \mathbf{A} and ϕ are solutions of the vector and scalar inhomogeneous Helmholtz equations. These equations can be solved by finding the Green's function $G(\mathbf{r})$ that is the solution of the scalar Helmholtz equation when the right hand side is a point source represented by a Dirac delta $\delta(\mathbf{r})$. Then, the superposition principle is used to get the solutions for \mathbf{A} and ϕ

$$\nabla^2 G(\mathbf{r}) + k^2 G(\mathbf{r}) = -\delta(\mathbf{r}) \quad (1.24a)$$

$$\mathbf{A}(\mathbf{r}) = \mu \iiint_{\mathbf{r}' \in \mathbb{R}^3} G(\mathbf{r} - \mathbf{r}') \mathbf{J}(\mathbf{r}') dV' \quad (1.24b)$$

$$\phi(\mathbf{r}) = \frac{1}{\varepsilon} \iiint_{\mathbf{r}' \in \mathbb{R}^3} G(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') dV'. \quad (1.24c)$$

To be solved uniquely, a boundary condition must be added to (1.24a). In free space, the boundary condition is imposed at infinity to allow outgoing waves only, so that in \mathbb{R}^3 with the $e^{i\omega t}$ harmonic convention, the Green's function G verifies the Sommerfeld's radiation condition

$$\lim_{|\mathbf{r}| \rightarrow +\infty} |\mathbf{r}| \left(\frac{\partial G(\mathbf{r})}{\partial |\mathbf{r}|} + ikG(\mathbf{r}) \right) = 0. \quad (1.25)$$

Then, the unique solution for the Green's function is [1]

$$G(\mathbf{r}) = \frac{e^{-ik|\mathbf{r}|}}{4\pi |\mathbf{r}|}. \quad (1.26)$$

So, the potentials are

$$\mathbf{A}(\mathbf{r}) = \mu \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi |\mathbf{r} - \mathbf{r}'|} \mathbf{J}(\mathbf{r}') dV' \quad (1.27a)$$

$$\phi(\mathbf{r}) = \frac{1}{\varepsilon} \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi |\mathbf{r} - \mathbf{r}'|} \rho(\mathbf{r}') dV' \quad (1.27b)$$

and the corresponding fields are

$$\mathbf{E}(\mathbf{r}) = -i\omega \mathbf{A}(\mathbf{r}) - \nabla \phi(\mathbf{r}) \quad (1.28a)$$

$$\mathbf{H}(\mathbf{r}) = \frac{1}{\mu} \nabla \times \mathbf{A}(\mathbf{r}). \quad (1.28b)$$

1.2.4 Potentials for the magnetic sources

While the magnetic sources are not actually physical, it is useful to generalize (or symmetrize) the Maxwell's equations by introducing

- ρ_m (Wb/m³) the magnetic charge density,
- \mathbf{M} (V/m²) the magnetic current density.

The generalized Maxwell's equations read

$$\nabla \cdot \mathbf{D}(\mathbf{r}) = \rho(\mathbf{r}) \quad \text{Maxwell-Gauss} \quad (1.29a)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}) = \rho_m(\mathbf{r}) \quad \text{Maxwell-Thomson} \quad (1.29b)$$

$$\nabla \times \mathbf{E}(\mathbf{r}) = -i\omega \mathbf{B}(\mathbf{r}) - \mathbf{M}(\mathbf{r}) \quad \text{Maxwell-Faraday} \quad (1.29c)$$

$$\nabla \times \mathbf{H}(\mathbf{r}) = i\omega \mathbf{D}(\mathbf{r}) + \mathbf{J}(\mathbf{r}) \quad \text{Maxwell-Ampère.} \quad (1.29d)$$

Assuming that there are no electric sources (i.e. $\rho = 0$ and $\mathbf{J} = \mathbf{0}$), it is possible to derive the potentials and the fields scattered by the magnetic sources ρ_m and \mathbf{M} . The results of the previous section are reused by taking advantage of the symmetry in Maxwell's equations and performing the adequate substitutions

$$\mathbf{E} \leftrightarrow \mathbf{H} \quad (1.30a)$$

$$\mathbf{B} \leftrightarrow -\mathbf{D} \quad (1.30b)$$

$$\rho \leftrightarrow -\rho_m \quad (1.30c)$$

$$\mathbf{J} \leftrightarrow -\mathbf{M} \quad (1.30d)$$

$$\varepsilon \leftrightarrow -\mu. \quad (1.30e)$$

It results in the following vector potential \mathbf{F} (in C/m) and scalar potential ϕ_m (in A)

$$\mathbf{F}(\mathbf{r}) = \varepsilon \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \mathbf{M}(\mathbf{r}') dV' \quad (1.31a)$$

$$\phi_m(\mathbf{r}) = \frac{1}{\mu} \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \rho_m(\mathbf{r}') dV'. \quad (1.31b)$$

The corresponding Lorenz gauge is

$$\nabla \cdot \mathbf{F}(\mathbf{r}) = -i\omega\mu\varepsilon\phi_m(\mathbf{r}) \quad (1.32)$$

and the corresponding fields are

$$\mathbf{E}(\mathbf{r}) = -\frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}) \quad (1.33a)$$

$$\mathbf{H}(\mathbf{r}) = -i\omega \mathbf{F}(\mathbf{r}) - \nabla \phi_m(\mathbf{r}). \quad (1.33b)$$

The equations (1.28) are the fields scattered by electric sources in the absence of magnetic sources. In contrast, (1.33) are the fields scattered by magnetic sources in the absence of electric sources. Using the superposition principle, the fields scattered by electric sources and magnetic sources are simply the sum of the fields in (1.28) and (1.33). Thus the fields scattered by ρ , ρ_m , \mathbf{J} and \mathbf{M} are

$$\mathbf{E}^{\text{sca}}(\mathbf{r}) = -i\omega \mathbf{A}(\mathbf{r}) - \nabla \phi(\mathbf{r}) - \frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}) \quad (1.34a)$$

$$\mathbf{H}^{\text{sca}}(\mathbf{r}) = -i\omega \mathbf{F}(\mathbf{r}) - \nabla \phi_m(\mathbf{r}) + \frac{1}{\mu} \nabla \times \mathbf{A}(\mathbf{r}). \quad (1.34b)$$

1. Background and notations

The Lorenz gauges (1.22) and (1.32) are used to substitute ϕ and ϕ_m , so that the scattered fields are expressed only in terms of the currents \mathbf{J} and \mathbf{M} via the vector potentials \mathbf{A} and \mathbf{F}

$$\mathbf{E}^{\text{sca}}(\mathbf{r}) = -i\omega\mathbf{A}(\mathbf{r}) + \frac{1}{i\omega\mu\varepsilon}\nabla\nabla\cdot\mathbf{A}(\mathbf{r}) - \frac{1}{\varepsilon}\nabla\times\mathbf{F}(\mathbf{r}) \quad (1.35a)$$

$$\mathbf{H}^{\text{sca}}(\mathbf{r}) = -i\omega\mathbf{F}(\mathbf{r}) + \frac{1}{i\omega\mu\varepsilon}\nabla\nabla\cdot\mathbf{F}(\mathbf{r}) + \frac{1}{\mu}\nabla\times\mathbf{A}(\mathbf{r}). \quad (1.35b)$$

1.2.5 Far field

1.2.5.1 Far field approximation

The equations for the scattered fields (1.34) are valid everywhere. However, it is usually useful to compute the fields far away from the sources, in which case some approximations can be done. Assume that the sources are located at \mathbf{r}' near the origin, and that the fields are computed at \mathbf{r} far from the origin. Thus $|\mathbf{r}'| \ll |\mathbf{r}|$. With this assumption it holds

$$|\mathbf{r} - \mathbf{r}'| \approx |\mathbf{r}| - \hat{\mathbf{r}} \cdot \mathbf{r}' \quad (1.36)$$

where $\hat{\mathbf{r}}$ denotes the direction of \mathbf{r} i.e.

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{|\mathbf{r}|}. \quad (1.37)$$

Then, the Green's function (1.26) can be approximated as

$$G(\mathbf{r} - \mathbf{r}') = \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \approx \frac{e^{-ik|\mathbf{r}|}}{4\pi|\mathbf{r}|} e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}'} \quad (1.38)$$

so that in the far field, the potentials (1.27a) and (1.31a) simplify to

$$\mathbf{A}(\mathbf{r}) = \mu \frac{e^{-ik|\mathbf{r}|}}{4\pi|\mathbf{r}|} \iiint_{\mathbf{r}' \in \mathbb{R}^3} e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}'} \mathbf{J}(\mathbf{r}') dV' \quad (1.39a)$$

$$\mathbf{F}(\mathbf{r}) = \varepsilon \frac{e^{-ik|\mathbf{r}|}}{4\pi|\mathbf{r}|} \iiint_{\mathbf{r}' \in \mathbb{R}^3} e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}'} \mathbf{M}(\mathbf{r}') dV'. \quad (1.39b)$$

1.2.5.2 Expression of the far fields

When $|\mathbf{r}| \rightarrow +\infty$, the radial parts of $i\omega\mathbf{A}(\mathbf{r})$ and $\nabla\phi(\mathbf{r})$ cancel each other and the non-radial part of the gradient of the scalar potential is negligible compared to the vector potential. Quantitatively, the potentials that are present in the expression of the electric field \mathbf{E}^{sca} (1.34a) have the following asymptotic behaviours when $|\mathbf{r}| \rightarrow$

$+\infty$

$$|\mathbf{A}(\mathbf{r})| = O\left(\frac{1}{|\mathbf{r}|}\right) \quad (1.40a)$$

$$|\hat{\mathbf{r}} \times \nabla \phi(\mathbf{r})| = O\left(\frac{1}{|\mathbf{r}|^2}\right) \quad (1.40b)$$

$$\hat{\mathbf{r}} \cdot (-i\omega \mathbf{A}(\mathbf{r}) - \nabla \phi(\mathbf{r})) = O\left(\frac{1}{|\mathbf{r}|^2}\right) \quad (1.40c)$$

$$|\nabla \times \mathbf{F}(\mathbf{r})| = O\left(\frac{1}{|\mathbf{r}|}\right). \quad (1.40d)$$

The potentials involved in the scattered magnetic field have analogous asymptotic behaviours. In practice, the scattered fields are still computed at a finite distance from the origin so that the radial components of the scattered fields must be removed from the vector potentials. Therefore, in the far field $\hat{\mathbf{r}} \cdot \mathbf{E}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = 0$ and $\hat{\mathbf{r}} \cdot \mathbf{H}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = 0$ are enforced by removing the radial components of the vector potentials

$$\mathbf{E}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = -i\omega (\mathbf{A}(\mathbf{r}) - (\mathbf{A}(\mathbf{r}) \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}) - \frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}) \quad (1.41a)$$

$$\mathbf{H}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = -i\omega (\mathbf{F}(\mathbf{r}) - (\mathbf{F}(\mathbf{r}) \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}) + \frac{1}{\mu} \nabla \times \mathbf{A}(\mathbf{r}). \quad (1.41b)$$

In addition to that, the scattered fields behave locally like a plane wave in the far field region. The fields of a plane wave have the form

$$\mathbf{E}_{\text{pw}}(\mathbf{r}) = E_0 e^{-i\mathbf{k} \cdot \mathbf{r}} \hat{\mathbf{p}} \quad (1.42a)$$

$$\mathbf{H}_{\text{pw}}(\mathbf{r}) = H_0 e^{-i\mathbf{k} \cdot \mathbf{r}} \hat{\mathbf{k}} \times \hat{\mathbf{p}} \quad (1.42b)$$

$$= \frac{1}{\eta} \hat{\mathbf{k}} \times \mathbf{E}_{\text{pw}}(\mathbf{r}) \quad (1.42c)$$

where

- E_0 and $H_0 = \frac{E_0}{\eta}$ are the peak amplitudes of the fields,
- \mathbf{k} is the wave vector such that $k = |\mathbf{k}|$ is the wave number and $\hat{\mathbf{k}}$ is the direction of propagation. Note that in the far field $\hat{\mathbf{k}} = \hat{\mathbf{r}}$,
- $\hat{\mathbf{p}}$ and $\hat{\mathbf{k}} \times \hat{\mathbf{p}}$ are the polarization of the electric field and the magnetic field,
- $\eta = \sqrt{\frac{\mu}{\varepsilon}}$ is the characteristic impedance of the medium.

The polarizations of the fields $\hat{\mathbf{p}}$ and $\hat{\mathbf{k}} \times \hat{\mathbf{p}}$ are orthogonal to the direction of propagation ($\hat{\mathbf{p}} \cdot \hat{\mathbf{k}} = 0$). In other words, the fields have no component along the direction of propagation, so $\hat{\mathbf{k}} \cdot \mathbf{E}_{\text{pw}}(\mathbf{r}) = 0$ and $\hat{\mathbf{k}} \cdot \mathbf{H}_{\text{pw}}(\mathbf{r}) = 0$.

Using the relation between \mathbf{E} and \mathbf{H} in a plane wave (1.42c), the curl of the vector potentials are simplified by considering two plane waves ($\mathbf{E}_{\text{FF1}}^{\text{sca}}, \mathbf{H}_{\text{FF1}}^{\text{sca}}$) scattered only

1. Background and notations

by \mathbf{M} (i.e. for which $\mathbf{A} = \mathbf{0}$ in (1.41)) and $(\mathbf{E}_{\text{FF2}}^{\text{sca}}, \mathbf{H}_{\text{FF2}}^{\text{sca}})$ scattered only by \mathbf{J} (i.e. for which $\mathbf{F} = \mathbf{0}$ in (1.41))

$$-\frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}) = \mathbf{E}_{\text{FF1}}^{\text{sca}}(\mathbf{r}) = -\eta \hat{\mathbf{r}} \times \mathbf{H}_{\text{FF1}}^{\text{sca}}(\mathbf{r}) = i\omega\eta \hat{\mathbf{r}} \times \mathbf{F}(\mathbf{r}) \quad (1.43a)$$

$$\frac{1}{\mu} \nabla \times \mathbf{A}(\mathbf{r}) = \mathbf{H}_{\text{FF2}}^{\text{sca}}(\mathbf{r}) = \frac{1}{\eta} \hat{\mathbf{r}} \times \mathbf{E}_{\text{FF2}}^{\text{sca}}(\mathbf{r}) = -\frac{i\omega}{\eta} \hat{\mathbf{r}} \times \mathbf{A}(\mathbf{r}). \quad (1.43b)$$

The expressions for the far fields finally read

$$\mathbf{E}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = -i\omega (\mathbf{A}(\mathbf{r}) - (\mathbf{A}(\mathbf{r}) \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}) + i\omega\eta \hat{\mathbf{r}} \times \mathbf{F}(\mathbf{r}) \quad (1.44a)$$

$$\mathbf{H}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = -i\omega (\mathbf{F}(\mathbf{r}) - (\mathbf{F}(\mathbf{r}) \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}) - \frac{i\omega}{\eta} \hat{\mathbf{r}} \times \mathbf{A}(\mathbf{r}) \quad (1.44b)$$

or equivalently using the wave number k ,

$$\mathbf{E}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = -ik\eta \frac{1}{\mu} (\mathbf{A}(\mathbf{r}) - (\mathbf{A}(\mathbf{r}) \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}) + ik\hat{\mathbf{r}} \times \frac{1}{\varepsilon} \mathbf{F}(\mathbf{r}) \quad (1.45a)$$

$$\mathbf{H}_{\text{FF}}^{\text{sca}}(\mathbf{r}) = -\frac{ik}{\eta} \frac{1}{\varepsilon} (\mathbf{F}(\mathbf{r}) - (\mathbf{F}(\mathbf{r}) \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}}) - ik\hat{\mathbf{r}} \times \frac{1}{\mu} \mathbf{A}(\mathbf{r}). \quad (1.45b)$$

1.2.5.3 Radar cross section

Let \mathbf{E}^{inc} be an incident electric field that induces some currents on a scatterer. These currents radiate the scattered electric field \mathbf{E}^{sca} . The Radar Cross Section (RCS) corresponds to the effective area (in m^2) of a scatterer that radiates back the electromagnetic fields in a given direction. It is defined as

$$\text{RCS}(\hat{\mathbf{r}}) = \lim_{|\mathbf{r}| \rightarrow +\infty} 4\pi |\mathbf{r}|^2 \frac{|\mathbf{E}^{\text{sca}}(\mathbf{r})|^2}{|\mathbf{E}^{\text{inc}}|^2} = 4\pi |\mathbf{r}|^2 \frac{|\mathbf{E}_{\text{FF}}^{\text{sca}}(\mathbf{r})|^2}{|\mathbf{E}^{\text{inc}}|^2}. \quad (1.46)$$

It must be noted that the far fields (1.44) vary in $O\left(\frac{1}{|\mathbf{r}|}\right)$ with the distance. Also, $|\mathbf{r}| |\mathbf{E}_{\text{FF}}^{\text{sca}}(\mathbf{r})|$ and $|\mathbf{r}| |\mathbf{H}_{\text{FF}}^{\text{sca}}(\mathbf{r})|$ depend only on the direction $\hat{\mathbf{r}}$. So the computation of the RCS with the far field is actually valid at any distance $|\mathbf{r}|$.

It is very convenient to display the RCS in logarithmic units, so that the RCS in dBsm (decibel relative to 1 m^2) is

$$\text{RCS}_{\text{dBsm}}(\hat{\mathbf{r}}) = 10 \log_{10} (\text{RCS}(\hat{\mathbf{r}})). \quad (1.47)$$

1.2.6 Boundary conditions on discontinuous interfaces

1.2.6.1 Integral forms of Maxwell's equations

When the medium of propagation is continuous in space, the fields are also continuous so the local Maxwell's equations (1.29) are valid. However, if there is an obstacle that introduces a discontinuity in the medium, the fields are also discontinuous across the

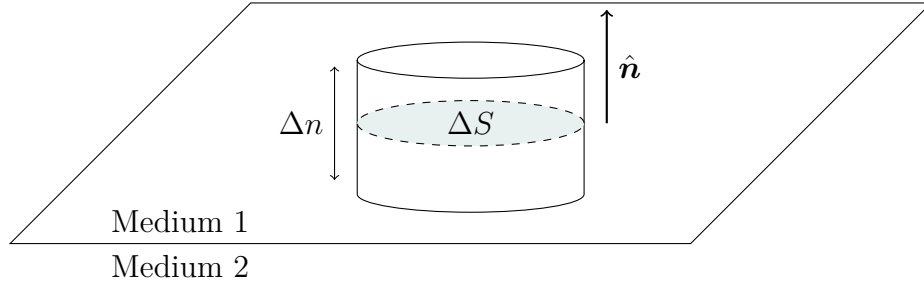


Figure 1.1: Cylinder across a discontinuous interface

interface. The local equations are not valid at discontinuities, so that the integral forms of Maxwell's equations are needed to find the relations between the fields in the two media. The integral forms of Maxwell's equations are always valid. In the frequency domain, for any volume V and surface S they read

$$\oiint_{\mathbf{r} \in \partial V} \mathbf{D}(\mathbf{r}) \cdot d\mathbf{S} = \iiint_{\mathbf{r} \in V} \rho(\mathbf{r}) dV \quad (1.48a)$$

$$\oiint_{\mathbf{r} \in \partial V} \mathbf{B}(\mathbf{r}) \cdot d\mathbf{S} = \iiint_{\mathbf{r} \in V} \rho_m(\mathbf{r}) dV \quad (1.48b)$$

$$\oint_{\mathbf{r} \in \partial S} \mathbf{E}(\mathbf{r}) \cdot d\mathbf{l} = -i\omega \iint_{\mathbf{r} \in S} \mathbf{B}(\mathbf{r}) \cdot d\mathbf{S} - \iint_{\mathbf{r} \in S} \mathbf{M}(\mathbf{r}) \cdot d\mathbf{S} \quad (1.48c)$$

$$\oint_{\mathbf{r} \in \partial S} \mathbf{H}(\mathbf{r}) \cdot d\mathbf{l} = i\omega \iint_{\mathbf{r} \in S} \mathbf{D}(\mathbf{r}) \cdot d\mathbf{S} + \iint_{\mathbf{r} \in S} \mathbf{J}(\mathbf{r}) \cdot d\mathbf{S}. \quad (1.48d)$$

In the following, the normal to the interface between the media 1 and 2 is noted \hat{n} and is pointing outside of the medium 2.

1.2.6.2 Boundary condition on the electric charges

The Maxwell-Gauss equation (1.48a) is used on a cylinder across the boundary as illustrated in the figure 1.1. The cylinder is assumed to be sufficiently small to consider locally that the interface is planar and that the fields are constant (\mathbf{D}_1 in medium 1 and \mathbf{D}_2 in medium 2). By letting the height of the cylinder $\Delta n \rightarrow 0$, the lateral contribution is vanishing so that

$$\oiint_{\mathbf{r} \in \partial V} \mathbf{D}(\mathbf{r}) \cdot d\mathbf{S} = \mathbf{D}_1 \cdot \hat{n} \Delta S - \mathbf{D}_2 \cdot \hat{n} \Delta S \quad (1.49a)$$

$$= (\mathbf{D}_1 - \mathbf{D}_2) \cdot \hat{n} \Delta S. \quad (1.49b)$$

In the limit of a very thin cylinder ($\Delta n \rightarrow 0$), all the charges lie on the boundary which results in a surface electric charge density ρ_s (in C/m²) that verifies

$$\iiint_{\mathbf{r} \in V} \rho(\mathbf{r}) dV = \iint_{\mathbf{r} \in S} \rho_s(\mathbf{r}) dS \quad (1.50a)$$

$$= \rho_s \Delta S. \quad (1.50b)$$

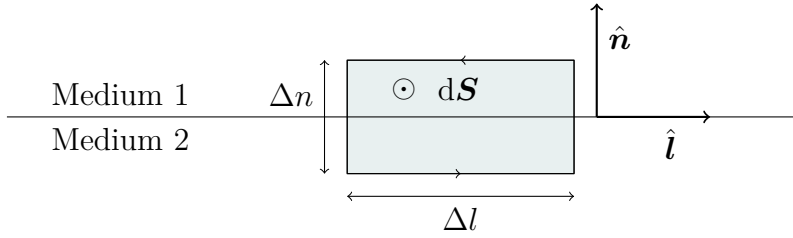


Figure 1.2: Rectangle across a discontinuous interface

Equating (1.49b) and (1.50b) results in the boundary condition

$$\hat{\mathbf{n}} \cdot (\mathbf{D}_1 - \mathbf{D}_2) = \rho_s. \quad (1.51)$$

1.2.6.3 Boundary condition on the electric currents

The Maxwell-Ampère equation (1.48d) is used on a rectangle across the boundary as illustrated in the figure 1.2. The rectangle is assumed to be sufficiently small to consider locally that the interface is planar and that the fields are constant (\mathbf{H}_1 in medium 1 and \mathbf{H}_2 in medium 2). By letting the height of the rectangle $\Delta n \rightarrow 0$, the lateral contribution is vanishing, so

$$\oint_{\mathbf{r} \in \partial S} \mathbf{H}(\mathbf{r}) \cdot d\mathbf{l} = -(\mathbf{H}_1 - \mathbf{H}_2) \cdot \hat{\mathbf{l}} \Delta l \quad (1.52a)$$

$$= (\hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2)) \cdot (\hat{\mathbf{l}} \times \hat{\mathbf{n}}) \Delta l. \quad (1.52b)$$

The second term cancels when $\Delta n \rightarrow 0$,

$$\lim_{\Delta n \rightarrow 0} \iint_{\mathbf{r} \in S} \mathbf{D}(\mathbf{r}) \cdot d\mathbf{S} = 0. \quad (1.53)$$

In the limit of a very thin rectangle ($\Delta n \rightarrow 0$), all the currents lie on the boundary which results in a surface electric current density \mathbf{J}_s (in A/m) that verifies

$$\iint_{\mathbf{r} \in S} \mathbf{J}(\mathbf{r}) \cdot d\mathbf{S} = \iint_{\mathbf{r} \in S} \mathbf{J}(\mathbf{r}) \cdot (\hat{\mathbf{l}} \times \hat{\mathbf{n}}) dS \quad (1.54a)$$

$$= \mathbf{J}_s \cdot (\hat{\mathbf{l}} \times \hat{\mathbf{n}}) \Delta l. \quad (1.54b)$$

Combining (1.52b) and (1.54b) results in

$$(\hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2)) \cdot (\hat{\mathbf{l}} \times \hat{\mathbf{n}}) = \mathbf{J}_s \cdot (\hat{\mathbf{l}} \times \hat{\mathbf{n}}). \quad (1.55)$$

As $\hat{\mathbf{l}} \times \hat{\mathbf{n}}$ is arbitrary, the above equation simplifies to

$$\hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_s. \quad (1.56)$$

1.2.6.4 Boundary conditions for the magnetic quantities

Thanks to the symmetry of the Maxwell's equations, it is possible to obtain the two remaining boundary conditions using the substitutions (1.30) in (1.51) and (1.56). The four boundary conditions are

$$\hat{\mathbf{n}} \cdot (\mathbf{D}_1 - \mathbf{D}_2) = \rho_s \quad (1.57a)$$

$$\hat{\mathbf{n}} \cdot (\mathbf{B}_1 - \mathbf{B}_2) = \rho_{ms} \quad (1.57b)$$

$$\hat{\mathbf{n}} \times (\mathbf{E}_1 - \mathbf{E}_2) = -\mathbf{M}_s \quad (1.57c)$$

$$\hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_s \quad (1.57d)$$

where notations similar to the previous ones have been used for the fields in the medium 1 (\mathbf{B}_1 and \mathbf{E}_1), the fields in the medium 2 (\mathbf{B}_2 and \mathbf{E}_2), the surface magnetic charge density ρ_{ms} (in Wb/m²) and the surface magnetic current density \mathbf{M}_s (in V/m). It should be noted that these boundary conditions were derived only from the Maxwell's equations so they don't depend on the material parameters of the media.

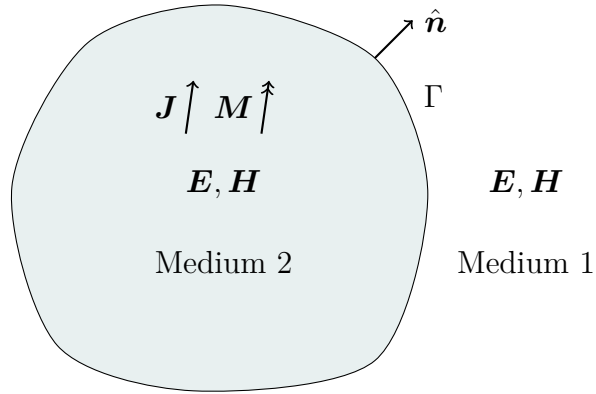
1.2.7 Surface integral equations

1.2.7.1 Surface equivalence principle

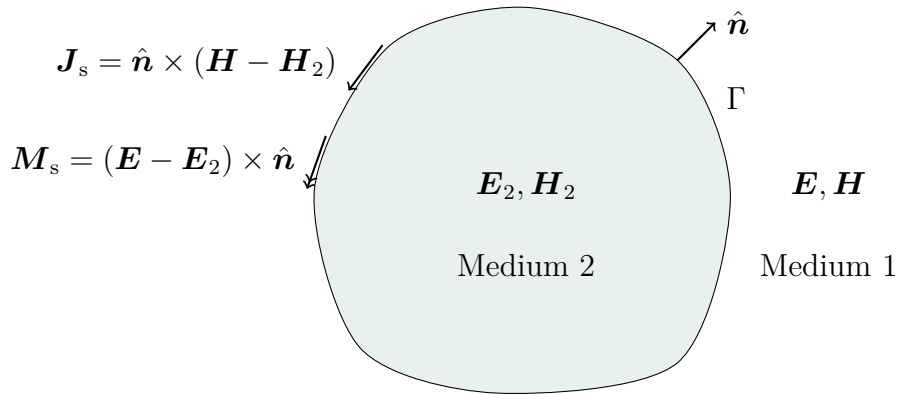
In a simple scattering problem, the background medium contains a scatterer whose surface is denoted Γ . Γ is the boundary between the background (medium 1) and the interior of the scatterer (medium 2). Given an electromagnetic excitation on the scatterer, the goal is to find the scattered fields \mathbf{E}^{sca} and \mathbf{H}^{sca} . This excitation can take various forms such as a voltage imposed on an antenna feed or electromagnetic fields impinging on a target. In this work, only this latter case is considered. The incident fields are noted \mathbf{E}^{inc} and \mathbf{H}^{inc} . They induce some currents \mathbf{J} and \mathbf{M} in the medium 2. These currents \mathbf{J} and \mathbf{M} induced by the incident fields act as sources for the scattered fields \mathbf{E}^{sca} and \mathbf{H}^{sca} as described in the sections 1.2.3 and 1.2.4. Therefore, the scattering problem is considered to be solved when the currents are found since \mathbf{E}^{sca} and \mathbf{H}^{sca} can be computed from \mathbf{J} and \mathbf{M} using (1.35).

However, care must be taken because the derivations done in 1.2.3 and 1.2.4 assume a free space, which is not true in the presence of a scatterer. In particular, the free space assumption is relevant in the determination of the Green's function (1.26). The Green's function can be determined for numerous canonical scenarios, but for a general geometry of Γ the Green's function is not accessible. Instead, the surface equivalence principle is needed.

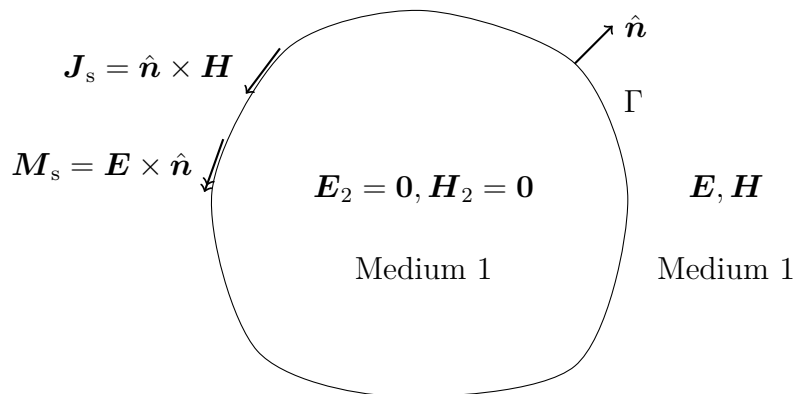
The surface equivalence principle transforms the scattering problem into an equivalent one that can be solved in free space. The initial configuration is illustrated in the figure 1.3a which corresponds to the original scattering problem where there are the true fields \mathbf{E} and \mathbf{H} everywhere in space. The key is to replace the fields \mathbf{E} and \mathbf{H} and the sources \mathbf{J} and \mathbf{M} in the medium 2 (inside Γ) by some other fields \mathbf{E}_2 and \mathbf{H}_2 , and by placing new source currents \mathbf{J}_s and \mathbf{M}_s on the surface Γ as illustrated in the figure 1.3b. These surface currents \mathbf{J}_s and \mathbf{M}_s are virtual but they are chosen to verify the boundary conditions derived in the section 1.2.6, so that the total fields



(a) Original problem



(b) Equivalent exterior problem



(c) Equivalent exterior problem with zero fields inside, so that the medium 2 can be replaced by the medium 1

Figure 1.3: Surface equivalence principle

\mathbf{E} and \mathbf{H} in the medium 1 (outside Γ) are left unchanged. In other words, outside Γ the fields scattered by \mathbf{J} and \mathbf{M} are the same as the fields scattered by \mathbf{J}_s and \mathbf{M}_s . In particular, the fields \mathbf{E}_2 and \mathbf{H}_2 inside Γ are chosen to be $\mathbf{0}$. In this case, the parameters of the medium 2 inside Γ can be replaced by the parameters of the medium 1 without affecting the fields inside that are $\mathbf{0}$ as illustrated in the figure 1.3c. By noting $\hat{\mathbf{n}}(\mathbf{r})$ the normal to Γ in \mathbf{r} , the currents are actually the tangential traces of the fields

$$\mathbf{J}_s(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}(\mathbf{r}) \quad (1.58a)$$

$$\mathbf{M}_s(\mathbf{r}) = \mathbf{E}(\mathbf{r}) \times \hat{\mathbf{n}}(\mathbf{r}). \quad (1.58b)$$

The problem has been transformed into a problem in free space as the medium 1 is the same everywhere in space. There is no obstacle but fictitious currents that radiate the same fields outside Γ . Henceforth, the potentials (1.27a) and (1.31a) can be used for the computation of the scattered fields (1.35). Again, these surface currents (1.58) are fictitious. The only situation where they match with the actual physical current is for Perfect Electric Conductors (PEC), in which case there is no magnetic current ($\mathbf{M}_s = \mathbf{E} \times \hat{\mathbf{n}} = \mathbf{0}$ on Γ).

The equivalent currents \mathbf{J}_s and \mathbf{M}_s are surfacic instead of volumic so the potentials (1.27a) and (1.31a) become

$$\mathbf{A}(\mathbf{r}) = \mu \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \mathbf{J}_s(\mathbf{r}') dS' \quad (1.59a)$$

$$\mathbf{F}(\mathbf{r}) = \varepsilon \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \mathbf{M}_s(\mathbf{r}') dS'. \quad (1.59b)$$

In the far field, the analogue of (1.39) using surface currents is

$$\mathbf{A}(\mathbf{r}) = \mu \frac{e^{-ik|\mathbf{r}|}}{4\pi|\mathbf{r}|} \iint_{\mathbf{r}' \in \Gamma} e^{ik\hat{\mathbf{r}} \cdot \mathbf{r}'} \mathbf{J}_s(\mathbf{r}') dS' \quad (1.60a)$$

$$\mathbf{F}(\mathbf{r}) = \varepsilon \frac{e^{-ik|\mathbf{r}|}}{4\pi|\mathbf{r}|} \iint_{\mathbf{r}' \in \Gamma} e^{ik\hat{\mathbf{r}} \cdot \mathbf{r}'} \mathbf{M}_s(\mathbf{r}') dS'. \quad (1.60b)$$

1.2.7.2 Surface integral equations

The total fields are written as a sum of the incident fields and the scattered fields

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{\text{inc}}(\mathbf{r}) + \mathbf{E}^{\text{sca}}(\mathbf{r}) \quad (1.61a)$$

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}^{\text{inc}}(\mathbf{r}) + \mathbf{H}^{\text{sca}}(\mathbf{r}) \quad (1.61b)$$

and inserted into the expression of the currents (1.58)

$$\mathbf{J}_s(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times (\mathbf{H}^{\text{inc}}(\mathbf{r}) + \mathbf{H}^{\text{sca}}(\mathbf{r})) \quad (1.62a)$$

$$\mathbf{M}_s(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r}) \times (\mathbf{E}^{\text{inc}}(\mathbf{r}) + \mathbf{E}^{\text{sca}}(\mathbf{r})). \quad (1.62b)$$

However, the expression of the scattered field that depends on the potentials (1.35) cannot be inserted directly in this previous expression (1.62) because the potentials

1. Background and notations

contain a singularity when they are evaluated on Γ . Let $\mathbf{r}_0 \in \Gamma$, the various quantity intervening in the scattered field (1.35) are investigated when $\mathbf{r} \rightarrow \mathbf{r}_0$ from the outside of Γ . The conclusions are directly given here but rigorous proofs can be found in [2]. For $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{E}^{\text{sca}}(\mathbf{r})$, the quantities $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{A}(\mathbf{r})$ and $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \nabla \nabla \cdot \mathbf{A}(\mathbf{r})$ behave well so they are simply written

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{A}(\mathbf{r}) = \mathbf{A}(\mathbf{r}_0) \quad (1.63a)$$

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \nabla \nabla \cdot \mathbf{A}(\mathbf{r}) = \nabla \nabla \cdot \mathbf{A}(\mathbf{r}_0) \quad (1.63b)$$

but $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \nabla \times \mathbf{F}(\mathbf{r})$ contains a singularity so that in \mathbf{r}_0 the integral is evaluated in terms of its principal value

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}) = \lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \iint_{\mathbf{r}' \in \Gamma} \nabla G(\mathbf{r} - \mathbf{r}') \times \mathbf{M}_s(\mathbf{r}') dS' \quad (1.64a)$$

$$= -\frac{\Omega(\mathbf{r}_0)}{4\pi} \hat{\mathbf{n}}(\mathbf{r}_0) \times \mathbf{M}_s(\mathbf{r}_0) \quad (1.64b)$$

$$+ \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla G(\mathbf{r}_0 - \mathbf{r}') \times \mathbf{M}_s(\mathbf{r}') dS' \quad (1.64c)$$

where $\Omega(\mathbf{r}_0)$ denotes the solid angle made by the exterior of Γ in \mathbf{r}_0 , which is 2π if Γ is smooth in \mathbf{r}_0 . So in all practical cases $\frac{\Omega(\mathbf{r}_0)}{4\pi} = \frac{1}{2}$, and in the following only this case is considered. Also, with some vector calculus manipulations and using the properties of the gradient of the Green's function, the following identity is obtained on Γ

$$\nabla \cdot \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \mathbf{J}_s(\mathbf{r}') dS' = \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{J}_s(\mathbf{r}') dS' \quad (1.65)$$

where $\nabla' \cdot \mathbf{J}_s(\mathbf{r}')$ denotes the divergence relative to \mathbf{r}' . In summary, on Γ ,

$$\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{sca}}(\mathbf{r}) = -ik\eta \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \mathbf{J}_s(\mathbf{r}') dS' \quad (1.66a)$$

$$+ \frac{\eta}{ik} \hat{\mathbf{n}}(\mathbf{r}) \times \nabla \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{J}_s(\mathbf{r}') dS' \quad (1.66b)$$

$$- \frac{1}{2} \mathbf{M}_s(\mathbf{r}) \quad (1.66c)$$

$$- \hat{\mathbf{n}}(\mathbf{r}) \times \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla G(\mathbf{r} - \mathbf{r}') \times \mathbf{M}_s(\mathbf{r}') dS' \quad (1.66d)$$

Similar results also hold for $\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \mathbf{H}^{\text{sca}}(\mathbf{r})$. So on Γ ,

$$\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}^{\text{sca}}(\mathbf{r}) = -\frac{ik}{\eta} \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \mathbf{M}_s(\mathbf{r}') dS' \quad (1.67a)$$

$$+ \frac{1}{ik\eta} \hat{\mathbf{n}}(\mathbf{r}) \times \nabla \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{M}_s(\mathbf{r}') dS' \quad (1.67b)$$

$$+ \frac{1}{2} \mathbf{J}_s(\mathbf{r}) \quad (1.67c)$$

$$+ \hat{\mathbf{n}}(\mathbf{r}) \times \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla G(\mathbf{r} - \mathbf{r}') \times \mathbf{J}_s(\mathbf{r}') dS'. \quad (1.67d)$$

1.2.7.3 Surface integral operators

At this point, it is useful to introduce surface integral operators to conveniently rewrite these equations. These operators depend on the wave number k (1.18), they act on a surface current (noted \mathbf{f} in the following) and they are evaluated on Γ . The operator \mathcal{T} has a singular part \mathcal{T}_s and an hypersingular part \mathcal{T}_h

$$\mathcal{T} = -ik\mathcal{T}_s + \frac{1}{ik}\mathcal{T}_h \quad (1.68a)$$

$$(\mathcal{T}_s \mathbf{f})(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \mathbf{f}(\mathbf{r}') dS' \quad (1.68b)$$

$$(\mathcal{T}_h \mathbf{f})(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \nabla \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \nabla' \cdot \mathbf{f}(\mathbf{r}') dS'. \quad (1.68c)$$

The operator \mathcal{K} is

$$(\mathcal{K} \mathbf{f})(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla \left(\frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \right) \times \mathbf{f}(\mathbf{r}') dS'. \quad (1.69)$$

The identity operator \mathcal{I} is

$$(\mathcal{I} \mathbf{f})(\mathbf{r}) = \mathbf{f}(\mathbf{r}). \quad (1.70)$$

All these operators are linear in \mathbf{f} . Using these operators, the tangential traces of the scattered fields on Γ are

$$\hat{\mathbf{n}} \times \mathbf{E}^{\text{sca}} = \eta \mathcal{T} \mathbf{J}_s - \left(\frac{\mathcal{I}}{2} + \mathcal{K} \right) \mathbf{M}_s \quad (1.71a)$$

$$\hat{\mathbf{n}} \times \mathbf{H}^{\text{sca}} = \frac{1}{\eta} \mathcal{T} \mathbf{M}_s + \left(\frac{\mathcal{I}}{2} + \mathcal{K} \right) \mathbf{J}_s. \quad (1.71b)$$

Inserting these expressions in (1.62) results in

$$\mathbf{M}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} - \eta \mathcal{T} \mathbf{J}_s + \left(\frac{\mathcal{I}}{2} + \mathcal{K} \right) \mathbf{M}_s \quad (1.72a)$$

$$\mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H}^{\text{inc}} + \frac{1}{\eta} \mathcal{T} \mathbf{M}_s + \left(\frac{\mathcal{I}}{2} + \mathcal{K} \right) \mathbf{J}_s \quad (1.72b)$$

or equivalently

$$\eta \mathcal{T} \mathbf{J}_s + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{M}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \quad (1.73a)$$

$$-\frac{1}{\eta} \mathcal{T} \mathbf{M}_s + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H}^{\text{inc}}. \quad (1.73b)$$

This first equation (1.73a) is the Electric Field Integral Equation (EFIE) and the second equation (1.73b) is the Magnetic Field Integral Equation (MFIE) for the exterior domain. These integral equations are conveniently written in a matrix form

$$\begin{pmatrix} \eta \mathcal{T} & \frac{\mathcal{I}}{2} - \mathcal{K} \\ \frac{\mathcal{I}}{2} - \mathcal{K} & -\frac{1}{\eta} \mathcal{T} \end{pmatrix} \begin{pmatrix} \mathbf{J}_s \\ \mathbf{M}_s \end{pmatrix} = \begin{pmatrix} -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \\ \hat{\mathbf{n}} \times \mathbf{H}^{\text{inc}} \end{pmatrix}. \quad (1.74)$$

1.2.7.4 Surface integral equations for the interior domain

Using a similar approach, it is possible to derive the EFIE and MFIE for the interior domain, by using the surface equivalence principle to set the zero fields outside of Γ . In this case, there are no incident fields. Let k' be the wave number and η' the characteristic impedance of the medium 2 (inside Γ), and let \mathcal{T}' and \mathcal{K}' be the corresponding operators (1.68a) and (1.69) with the wave number k' (note that the normal $\hat{\mathbf{n}}$ in the definition of \mathcal{T}' and \mathcal{K}' is still the same as \mathcal{T} and \mathcal{K} i.e. pointing outside Γ). Then, the EFIE and MFIE for the interior domain are

$$-\eta' \mathcal{T}' \mathbf{J}_s + \left(\frac{\mathcal{I}}{2} + \mathcal{K}' \right) \mathbf{M}_s = \mathbf{0} \quad (1.75a)$$

$$\frac{1}{\eta'} \mathcal{T}' \mathbf{M}_s + \left(\frac{\mathcal{I}}{2} + \mathcal{K}' \right) \mathbf{J}_s = \mathbf{0}. \quad (1.75b)$$

In a matrix form they read

$$\begin{pmatrix} -\eta' \mathcal{T}' & \frac{\mathcal{I}}{2} + \mathcal{K}' \\ \frac{\mathcal{I}}{2} + \mathcal{K}' & \frac{1}{\eta'} \mathcal{T}' \end{pmatrix} \begin{pmatrix} \mathbf{J}_s \\ \mathbf{M}_s \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}. \quad (1.76)$$

1.3 Boundary element method

1.3.1 Method of moments

The Boundary Element Method (BEM) is the technique that applies the Method of Moments (MoM) on a boundary value problem formulated as an integral equation. In the case of the surface integral equations derived in the section 1.2.7, the values at the boundary are the tangential traces of the electric and magnetic fields i.e. the fictitious surface currents \mathbf{J}_s and \mathbf{M}_s (1.58).

Let $\mathcal{L} : X \rightarrow Y$ be a linear operator acting on an unknown function $\mathbf{x} \in X$, and let $\mathbf{y} \in Y$ be a known excitation function. In general \mathbf{x} and \mathbf{y} are not necessarily in the same space of functions. For example, $X, Y \subset \{\mathbf{f} : \Gamma \rightarrow \mathbb{C}^3\}$ in the case of

the surface integral equations derived in the section 1.2.7. The MoM transforms the problem

$$\mathcal{L}\mathbf{x} = \mathbf{y} \quad (1.77)$$

into a linear system that can be solved

$$\mathbf{L}\mathbf{x} = \mathbf{y} \quad (1.78)$$

where \mathbf{L} is a matrix that corresponds to the discretization of the operator \mathcal{L} , whereas \mathbf{x} and \mathbf{y} are vectors that correspond the discretizations of \mathbf{x} and \mathbf{y} .

The unknown \mathbf{x} is approximated as a linear combination of N_s source basis functions $(\mathbf{s}_n)_{n=1}^{N_s}$ ($\mathbf{s}_n : \Gamma \rightarrow \mathbb{R}^3$). The coefficients $(x_n)_{n=1}^{N_s}$ associated to each basis function are stored in a vector $\mathbf{x} \in \mathbb{C}^{N_s}$, so that for all $\mathbf{r} \in \Gamma$

$$\mathbf{x}(\mathbf{r}) \approx \sum_{n=1}^{N_s} x_n \mathbf{s}_n(\mathbf{r}). \quad (1.79)$$

By linearity (1.77) becomes

$$\sum_{n=1}^{N_s} x_n \mathcal{L}\mathbf{s}_n \approx \mathbf{y}. \quad (1.80)$$

The error of approximation $\mathbf{y} - \sum_{n=1}^{N_s} x_n \mathcal{L}\mathbf{s}_n$ is the so-called residual that is to be minimized. This residual is tested by a set of N_s testing basis functions $(\mathbf{t}_m)_{m=1}^{N_s}$ ($\mathbf{t}_m : \Gamma \rightarrow \mathbb{R}^3$). The testing procedure consists in computing the inner product between \mathbf{t}_m and the residual for each $m \in [1, N_s]$. The inner product between two functions \mathbf{a} and $\mathbf{b} : \Gamma \rightarrow \mathbb{C}^3$ is defined as

$$\langle \mathbf{a}, \mathbf{b} \rangle = \iint_{\mathbf{r} \in \Gamma} \overline{\mathbf{a}(\mathbf{r})} \cdot \mathbf{b}(\mathbf{r}) \, dS. \quad (1.81)$$

The overline denotes a complex conjugation. For all $m \in [1, N_s]$, imposing that the testing of the residual is 0 results in

$$\left\langle \mathbf{t}_m, \mathbf{y} - \sum_{n=1}^{N_s} x_n \mathcal{L}\mathbf{s}_n \right\rangle = 0 \iff \sum_{n=1}^{N_s} x_n \langle \mathbf{t}_m, \mathcal{L}\mathbf{s}_n \rangle = \langle \mathbf{t}_m, \mathbf{y} \rangle \quad (1.82a)$$

$$\iff \sum_{n=1}^{N_s} \mathbf{L}_{mn} x_n = y_m \quad (1.82b)$$

$$\iff \mathbf{L}\mathbf{x} = \mathbf{y} \quad (1.82c)$$

where

$$\mathbf{L}_{mn} = \langle \mathbf{t}_m, \mathcal{L}\mathbf{s}_n \rangle \quad (1.83a)$$

$$y_m = \langle \mathbf{t}_m, \mathbf{y} \rangle. \quad (1.83b)$$

This linear system $\mathbf{L}\mathbf{x} = \mathbf{y}$ is then solved for \mathbf{x} . Finally, when the array of coefficients \mathbf{x} is known, the unknown \mathbf{x} is retrieved using the linear combination of source basis functions (1.79).

1.3.2 Discretization of the composition of operators

Let $\mathcal{L} : A \rightarrow B$ and $\mathcal{M} : C \rightarrow D$ be two operators ($B \subset C$), and for all $n \in [1, N_s]$ $\mathbf{a}_n \in A$ and $\mathbf{c}_n \in C$ are sources basis functions and $\mathbf{b}_n \in B^*$ and $\mathbf{d}_n \in D^*$ are testing basis functions (where $*$ denotes the dual space). The MoM discretizations (1.83a) of \mathcal{L} and \mathcal{M} are the matrices \mathbf{L} and \mathbf{M} whose elements are

$$\mathbf{L}_{mn} = \langle \mathbf{b}_m, \mathcal{L}\mathbf{a}_n \rangle \quad (1.84a)$$

$$\mathbf{M}_{mn} = \langle \mathbf{d}_m, \mathcal{M}\mathbf{c}_n \rangle. \quad (1.84b)$$

The discretization of $\mathcal{M}\mathcal{L}$ can be obtained from \mathbf{L} and \mathbf{M} by considering the functions $\mathbf{x} \in A$, $\mathbf{y} \in B$ and $\mathbf{z} \in D$ and their respective discretizations \mathbf{x} , \mathbf{y} and \mathbf{z} such that

$$\mathbf{x} = \sum_{n=1}^{N_s} x_n \mathbf{a}_n \quad (1.85a)$$

$$\mathbf{y} = \mathcal{L}\mathbf{x} = \sum_{n=1}^{N_s} y_n \mathbf{c}_n \quad (1.85b)$$

$$\mathbf{z} = \mathcal{M}\mathbf{y} \quad (1.85c)$$

$$\mathbf{z} = \mathbf{M}\mathbf{y}. \quad (1.85d)$$

Rewriting (1.85b) as $\mathcal{I}\mathbf{y} = \mathcal{L}\mathbf{x}$ and testing it with \mathbf{b}_m for all $m \in [1, N_s]$ results in

$$\mathbf{G}_m \mathbf{y} = \mathbf{L}\mathbf{x} \quad (1.86)$$

where \mathbf{G}_m is the so-called mix-Gram Matrix which is the discretization of the identity operator

$$[\mathbf{G}_m]_{mn} = \langle \mathbf{b}_m, \mathcal{I}\mathbf{c}_n \rangle = \langle \mathbf{b}_m, \mathbf{c}_n \rangle. \quad (1.87)$$

If the source and testing are the same i.e. $\mathbf{b}_m = \mathbf{c}_m$ for all $m \in [1, N_s]$ then it is a so-called Gram matrix noted \mathbf{G}

$$[\mathbf{G}]_{mn} = \langle \mathbf{b}_m, \mathbf{b}_n \rangle. \quad (1.88)$$

Assuming that the matrix \mathbf{G}_m is invertible, combining (1.85d) with (1.86) results in

$$\mathbf{z} = \mathbf{M}\mathbf{G}_m^{-1}\mathbf{L}\mathbf{x} \quad (1.89)$$

which is the discretization of the equation $\mathbf{z} = \mathcal{M}\mathcal{L}\mathbf{x}$. In other words, $\mathcal{M}\mathcal{L}$ is discretized by $\mathbf{M}\mathbf{G}_m^{-1}\mathbf{L}$.

This composition of operators is particularly relevant in the Calderón preconditioning (see section 1.5.3.1).

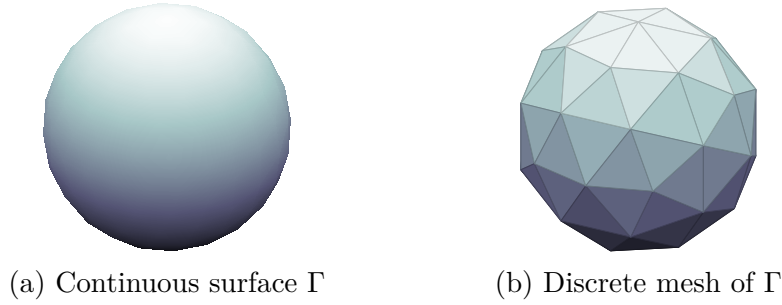


Figure 1.4: Discretization of a surface by a triangular mesh

1.3.3 Basis functions

In order to build a set of basis functions on a surface Γ , this surface must be discretized. As illustrated in the figure 1.4, the surface is approximated by a triangular mesh containing N_v vertices, N_f flat triangular faces, and N_s edges. In addition, the surface Γ has N_h handles. If $N_h = 0$ the geometry is simply connected (e.g. a sphere), otherwise it is multiply connected (e.g. $N_h = 1$ on a torus). The surface Γ is the boundary between the interior and the exterior domain such that it is oriented with its normal $\hat{\mathbf{n}}$ pointing toward the exterior domain. Furthermore, Γ does not contain holes and has no boundary i.e. $\partial\Gamma = \emptyset$. In other words, sheet scatterers, open structures and junctions are not considered in this work.

Each vertex is given an index $n \in [1, N_v]$ and noted \mathbf{v}_n . Each triangle face is given an index $n \in [1, N_f]$ and noted T_n . Each edge is given an index $n \in [1, N_s]$, noted e_n , and is given an arbitrary orientation from a vertex noted \mathbf{e}_n^- to a vertex \mathbf{e}_n^+ ($\mathbf{e}_n^-, \mathbf{e}_n^+ \in \{\mathbf{v}_m \text{ with } m \in [1, N_v]\}$). The average edge length is noted h .

While they are technically different, no distinction is made between the surface Γ and its triangular mesh approximation that is also called Γ in the following.

Local basis functions have their supports (i.e. where they are different than 0) that does not span on the entire mesh. The local basis functions defined in the following are based either on a single vertex, on a single edge, or on a single face. As opposed to the local basis functions there are also basis functions that are based on the entire geometry of the mesh (e.g. the global loops that are based on handles).

1.3.3.1 Rao-Wilton-Glisson basis functions

The Rao-Wilton-Glisson (RWG) [3] basis functions \mathbf{f}_n are based on the edges ($n \in [1, N_s]$). The support of the RWG basis function \mathbf{f}_n is made of the two faces T_n^- and T_n^+ that share the edge e_n . Using the notations of the figure 1.5, for each $n \in [1, N_s]$

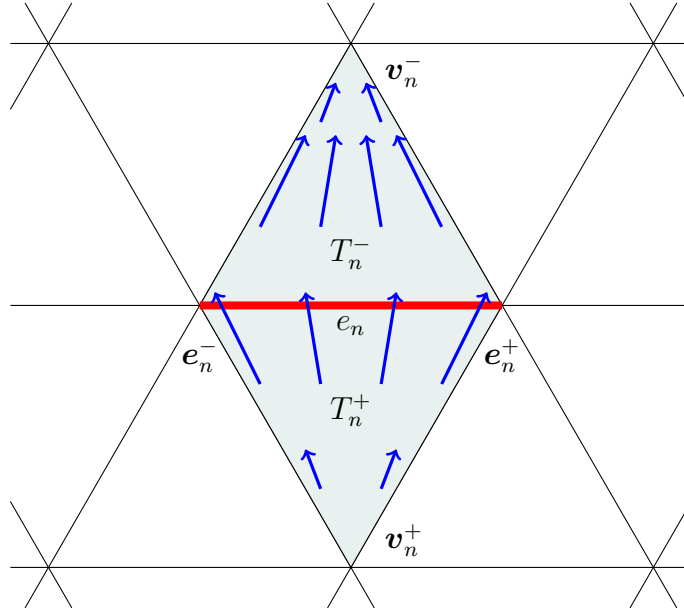


Figure 1.5: Rao-Wilton-Glisson (RWG) basis function \mathbf{f}_n

the RWG basis function \mathbf{f}_n is defined as

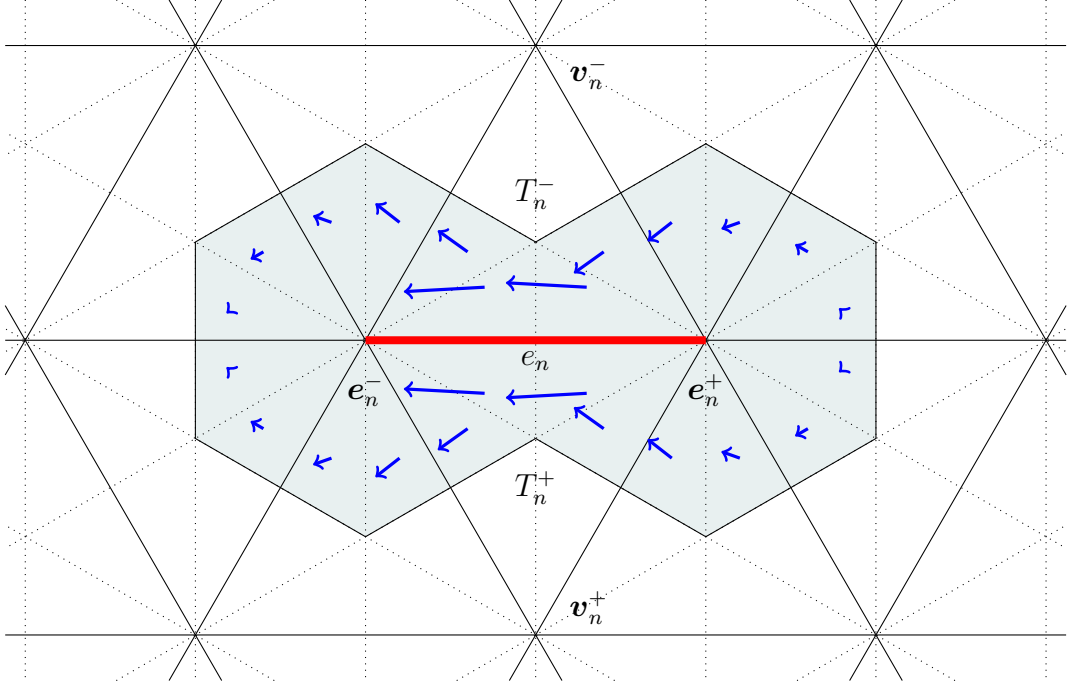
$$\mathbf{f}_n(\mathbf{r}) = \begin{cases} \frac{\mathbf{r} - \mathbf{v}_n^+}{2A_n^+} & \text{if } \mathbf{r} \in T_n^+ \\ \frac{\mathbf{v}_n^- - \mathbf{r}}{2A_n^-} & \text{if } \mathbf{r} \in T_n^- \\ \mathbf{0} & \text{otherwise} \end{cases} \quad (1.90)$$

where A_n^\pm is the area of T_n^\pm .

In the literature, they are sometimes called Raviart-Thomas basis functions. Also, they are often scaled by the length of the edge they are based on (including in the original paper [3]). The reason why they are not scaled by the edge length in (1.90) is that it simplifies the expression of the BC, loop and star basis functions that are introduced later.

When the RWG basis functions discretize a current, it flows across the edge from \mathbf{v}_n^+ to \mathbf{v}_n^- . The value of the function at various positions on the support as been represented in the figure 1.5. The divergence of the RWG basis functions $\nabla \cdot \mathbf{f}_n$ discretizes the charges. It is actually piece-wise constant patches

$$\nabla \cdot \mathbf{f}_n(\mathbf{r}) = \begin{cases} \frac{1}{A_n^+} & \text{if } \mathbf{r} \in T_n^+ \\ -\frac{1}{A_n^-} & \text{if } \mathbf{r} \in T_n^- \\ 0 & \text{otherwise.} \end{cases} \quad (1.91)$$


 Figure 1.6: Buffa-Christiansen (BC) basis function \mathbf{g}_n

1.3.3.2 Buffa-Christiansen basis functions

The Buffa-Christiansen (BC) [4] basis functions \mathbf{g}_n are also based on edges ($n \in [1, N_s]$). However, their definition requires the construction of a barycentric refinement of the mesh. To build the barycentric refinement of a mesh, a new vertex is spawned at the center of each edge and each face, so that each face in the primal mesh is subdivided into 6 faces in the barycentric refinement. In the illustration of the BC basis function (figure 1.6), the primal mesh is in solid line and the barycentric mesh is in dotted line.

The BC basis functions are actually linear combinations of RWG basis functions built on the barycentric mesh. Let N_+ be the number of edges that share the vertex \mathbf{e}_n^+ , and N_- be the number of edges that share the vertex \mathbf{e}_n^- in the primal mesh ($N_+ = N_- = 6$ in the figure 1.6). So in the barycentric mesh there are respectively $2N_+$ and $2N_-$ edges sharing \mathbf{e}_n^+ and \mathbf{e}_n^- . Let \mathbf{f}_k^{b+} ($k \in [0, 2N_+ - 1]$) be the RWG basis functions of the barycentric mesh whose edges share the central vertex \mathbf{e}_n^+ , ordered counterclockwise around the central vertex, starting from \mathbf{f}_0^{b+} on the reference edge e_n . Similarly, \mathbf{f}_ℓ^{b-} ($\ell \in [0, 2N_- - 1]$) are the RWG basis functions of the barycentric mesh sharing the central vertex \mathbf{e}_n^- , ordered counterclockwise starting from \mathbf{f}_0^{b-} on the reference edge e_n . Also, in the center of the BC function there are two barycentric RWG functions whose respective \mathbf{v}^+ and \mathbf{v}^- coincide with \mathbf{e}_n^+ and \mathbf{e}_n^- : the one on T_n^- is noted \mathbf{f}_-^{b0} and the one on T_n^+ is noted \mathbf{f}_+^{b0} . Then the BC basis function \mathbf{g}_n is

$$\mathbf{g}_n = \frac{1}{2} (s_-^{b0} \mathbf{f}_-^{b0} - s_+^{b0} \mathbf{f}_+^{b0}) + \sum_{k=1}^{2N_+-1} \frac{N_+ - k}{2N_+} s_k^{b+} \mathbf{f}_k^{b+} - \sum_{\ell=1}^{2N_- - 1} \frac{N_- - \ell}{2N_-} s_\ell^{b-} \mathbf{f}_\ell^{b-} \quad (1.92)$$

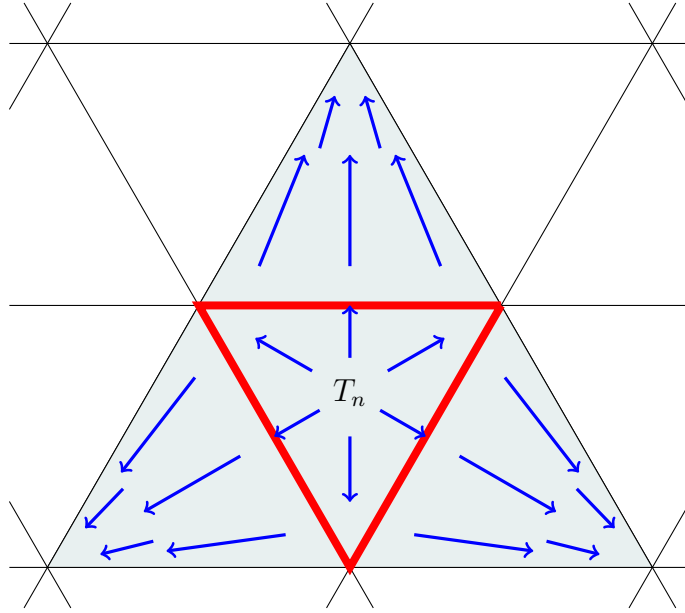


Figure 1.7: Star basis function σ_n

The sign coefficient s_x^{by} are $+1$ or -1 depending on the orientation of the corresponding \mathbf{f}_x^{by} function ($x, y \in \{-, +, 0, 1, 2, \dots\}$). $s_x^{\text{by}} = +1$ if the barycentric RWG function \mathbf{f}_x^{by} is oriented such that its vertex e^- is on the reference edge e_n of the BC function and its vertex e^+ is on the boundary of the support of the BC function. Otherwise, $s_x^{\text{by}} = -1$. When the BC basis functions discretize a current, it flows along the edge e_n from e_n^+ to e_n^- as it is illustrated in the figure 1.6.

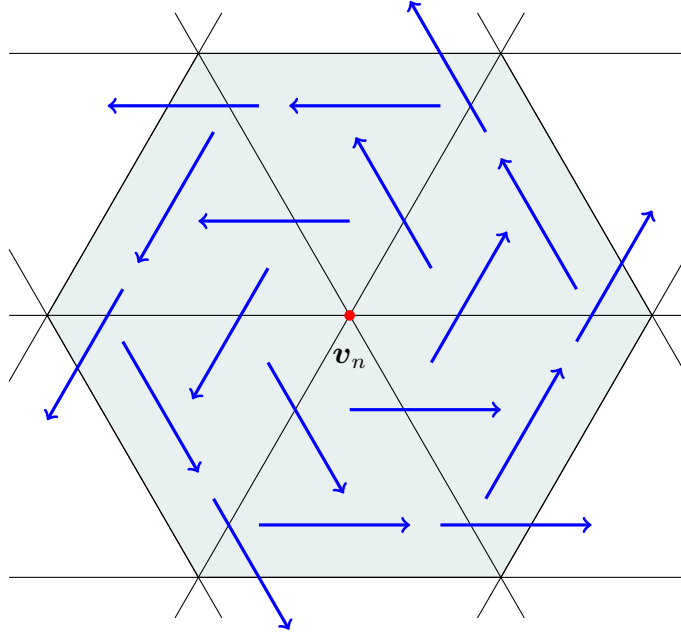
1.3.3.3 Star basis functions

The star basis functions σ_n [5] are based on the faces of the mesh ($n \in [1, N_f]$). σ_n is a linear combination of the 3 RWG basis functions whose reference edges are on the boundary of the reference face T_n of the star function. Or equivalently, the star function σ_n is the combination of the 3 RWG functions whose supports contain the reference faces T_n . For each $n \in [1, N_f]$, the star function σ_n is

$$\sigma_n = \sum_{m=1}^{N_s} \Sigma_{mn} \mathbf{f}_m \quad (1.93)$$

where Σ_{mn} are the coefficients of the Star-to-RWG basis transformation matrix $\Sigma \in \mathbb{R}^{N_s \times N_f}$. As explained above, there are only 3 non-zero terms in the sum (1.93) so Σ is sparse, and for $m \in [1, N_s], n \in [1, N_f]$ its coefficients are

$$\Sigma_{mn} = \begin{cases} +1 & \text{if the edge } e_m \text{ is on the boundary of the face } T_n \text{ clockwise} \\ -1 & \text{if the edge } e_m \text{ is on the boundary of the face } T_n \text{ counterclockwise} \\ 0 & \text{otherwise,} \end{cases} \quad (1.94)$$


 Figure 1.8: Local loop basis function λ_n

or equivalently

$$\mathbf{\Sigma}_{mn} = \begin{cases} +1 & \text{if } T_m^+ = T_n \\ -1 & \text{if } T_m^- = T_n \\ 0 & \text{otherwise.} \end{cases} \quad (1.95)$$

For the discretization of the current using the star functions, the current flows from the inside of the reference face to the outside as it is illustrated in the figure 1.7.

1.3.3.4 Loop basis functions

The local loop basis functions λ_n [5] are based on the vertices of the mesh ($n \in [1, N_v]$). They are a linear combination of the RWG functions whose reference edges contain the reference vertex \mathbf{v}_n of the loop functions. For each $n \in [1, N_v]$, the loop function λ_n is

$$\lambda_n = \sum_{m=1}^{N_s} \mathbf{\Lambda}_{mn} \mathbf{f}_m \quad (1.96)$$

where $\mathbf{\Lambda}_{mn}$ are the coefficients of the Loop-to-RWG basis transformation matrix $\mathbf{\Lambda} \in \mathbb{R}^{N_s \times N_v}$. As explained above, the number of non-zero terms in the sum (1.96) is equal to the number of vertices adjacent to \mathbf{v}_n so $\mathbf{\Lambda}$ is sparse, and for $m \in [1, N_s]$, $n \in [1, N_v]$ its coefficients are

$$\mathbf{\Lambda}_{mn} = \begin{cases} +1 & \text{if the edge } e_m \text{ leaves the vertex } \mathbf{v}_n \\ -1 & \text{if the edge } e_m \text{ arrives at the vertex } \mathbf{v}_n \\ 0 & \text{otherwise,} \end{cases} \quad (1.97)$$

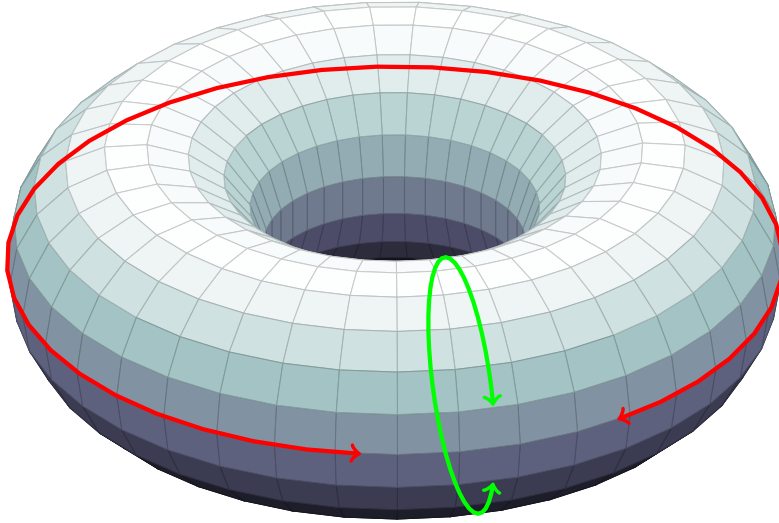


Figure 1.9: Global loops on a torus, toroidal (red) and poloidal (green)

or equivalently

$$\Lambda_{mn} = \begin{cases} +1 & \text{if } \mathbf{e}_m^- = \mathbf{v}_n \\ -1 & \text{if } \mathbf{e}_m^+ = \mathbf{v}_n \\ 0 & \text{otherwise.} \end{cases} \quad (1.98)$$

For the discretization of the current using the local loop functions, the current flows around the reference vertex counterclockwise as it is illustrated in the figure 1.8.

It is clear from the construction of the local loops that the divergences of two RWG functions (1.91) cancel each other on each triangle of their supports. So for all $n \in [1, N_v]$ and $\mathbf{r} \in \Gamma$

$$\nabla \cdot \boldsymbol{\lambda}_n(\mathbf{r}) = 0. \quad (1.99)$$

So the loop functions are solenoidal. In particular, there are scalar functions noted $P_n : \Gamma \rightarrow \mathbb{R}$ ($n \in [1, N_v]$) such that

$$\boldsymbol{\lambda}_n(\mathbf{r}) = \hat{\mathbf{n}}(\mathbf{r}) \times \nabla P_n(\mathbf{r}). \quad (1.100)$$

These functions P_n are the so-called pyramids, they have the same supports as the local loop functions and they are defined as

$$P_n(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} = \mathbf{v}_n \\ 0 & \text{if } \mathbf{r} \text{ is on the boundary of the support or not on the support} \\ \text{linear} & \text{elsewhere on the support.} \end{cases} \quad (1.101)$$

1.3.3.5 Global loops

As opposed to the local loops there are also the global loops $\boldsymbol{\eta}_n$ that exist only on multiply connected geometries ($n \in [1, 2N_h]$). For each handle there are 2 global

loops that are associated to it. These two global loops have been represented on the figure 1.9 for a torus. For each $n \in [1, 2N_h]$, the global loop function $\boldsymbol{\eta}_n$ is

$$\boldsymbol{\eta}_n = \sum_{m=1}^{N_s} \mathbf{H}_{mn} \mathbf{f}_m \quad (1.102)$$

where \mathbf{H}_{mn} are the coefficients of the global loop-to-RWG basis transformation matrix $\mathbf{H} \in \mathbb{R}^{N_s \times 2N_h}$. Similarly to the local loops, they are also solenoidal ($\nabla \cdot \boldsymbol{\eta}_n = 0$). Finding these global loops is computationally expensive in general so \mathbf{H} is not explicit.

1.3.4 Discretization of the surface integral operators

In the context of the Method of Moments (MoM), the Galerkin method refers to the case where the testing functions are the same as the source basis functions. The Petrov-Galerkin theory provides the mathematical foundations to ensure that the MoM converge to the analytic solution under certain conditions. In other words, the choice of source and testing basis functions should not be done thoughtlessly. The conforming discretizations of the operators introduced in the section 1.2.7.3 are given in the following.

The operator \mathcal{T} defined in (1.68a), is discretized conformingly using the RWG basis functions \mathbf{f}_n as sources, and the rotated RWG basis functions $\hat{\mathbf{n}} \times \mathbf{f}_m$ as testing ($m, n \in [1, N_s]$), so that the MoM matrix $\mathbf{T} \in \mathbb{C}^{N_s \times N_s}$ that discretizes \mathcal{T} with RWG basis functions is

$$\mathbf{T} = -ik\mathbf{T}_s + \frac{1}{ik}\mathbf{T}_h \quad (1.103)$$

where \mathbf{T}_s and \mathbf{T}_h discretize \mathcal{T}_s and \mathcal{T}_h , and are defined as

$$[\mathbf{T}_s]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_s \mathbf{f}_n \rangle \quad (1.104a)$$

$$[\mathbf{T}_h]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_h \mathbf{f}_n \rangle. \quad (1.104b)$$

For example, the Perfect Electric Conductor (PEC)-EFIE that reads

$$\eta \mathcal{T} \mathbf{J}_s(\mathbf{r}) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}) \quad (1.105)$$

is discretized as

$$\eta \mathbf{T} \mathbf{J} = \mathbf{E} \quad (1.106)$$

where the RWG basis functions \mathbf{f}_n are used as sources for the electric current (see (1.79))

$$\mathbf{J}_s(\mathbf{r}) \approx \sum_{n=1}^{N_s} \mathbf{J}_n \mathbf{f}_n(\mathbf{r}) \quad (1.107)$$

and the Right Hand Side (RHS) is tested with $\hat{\mathbf{n}} \times \mathbf{f}_m$ ($m \in [1, N_s]$) as

$$\mathbf{E}_m = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle = -\langle \mathbf{f}_m, \mathbf{E}^{\text{inc}} \rangle. \quad (1.108)$$

1. Background and notations

Similarly, the operator \mathcal{T} is also discretized with BC basis functions where \mathbf{g}_n are sources and $\hat{\mathbf{n}} \times \mathbf{g}_m$ are testing basis functions ($m, n \in [1, N_s]$)

$$\mathbb{T} = -ik\mathbb{T}_s + \frac{1}{ik}\mathbb{T}_h \quad (1.109a)$$

$$[\mathbb{T}]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{g}_m, \mathcal{T}\mathbf{g}_n \rangle \quad (1.109b)$$

$$[\mathbb{T}_s]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{g}_m, \mathcal{T}_s\mathbf{g}_n \rangle \quad (1.109c)$$

$$[\mathbb{T}_h]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{g}_m, \mathcal{T}_h\mathbf{g}_n \rangle. \quad (1.109d)$$

The operator \mathcal{K} is discretized conformingly using a mixed discretization. If the source are BC functions \mathbf{g}_n then the testing functions should be rotated RWG functions $\hat{\mathbf{n}} \times \mathbf{f}_m$. Then, the MoM discretization of the operator \mathcal{K} is the matrix \mathbf{K} defined by

$$[\mathbf{K}]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{K}\mathbf{g}_n \rangle. \quad (1.110)$$

The Galerkin testing of the operator \mathcal{K} with RWG basis functions (\mathbf{f}_n) as it has been done classically ($\langle \hat{\mathbf{n}} \times \mathbf{f}_n, \mathcal{K}\mathbf{f}_m \rangle$) is not conforming and results in an inaccurate discretization [6].

The identity operator \mathcal{I} can be discretized with the same source and testing basis functions, in which case it is a so-called Gram matrix (1.88). The Gram matrix for RWG basis functions (\mathbf{f}_n) is noted $\mathbf{G} \in \mathbb{R}^{N_s \times N_s}$

$$\mathbf{G}_{mn} = \langle \mathbf{f}_m, \mathbf{f}_n \rangle. \quad (1.111)$$

The identity operator can also be discretized with testing basis functions different than the sources, in this case it is the so-called mix-Gram matrix (1.87). It was already introduced in the section 1.3.2 as its inverse is used to compose operators. The mix-Gram matrix with rotated RWG functions ($\hat{\mathbf{n}} \times \mathbf{f}_m$) in testing and BC functions (\mathbf{g}_n) in sources is noted $\mathbf{G}_m \in \mathbb{R}^{N_s \times N_s}$

$$[\mathbf{G}_m]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathbf{g}_n \rangle. \quad (1.112)$$

Note that the mix-Gram matrix with RWG (\mathbf{f}_n) and rotated RWG ($\hat{\mathbf{n}} \times \mathbf{f}_m$) basis functions (whose elements are $\langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathbf{f}_n \rangle$) is singular, but \mathbf{G}_m is not. This is why the discretization of \mathcal{T}^2 uses a mixed discretization $\mathbb{T}\mathbf{G}_m^{-1}\mathbb{T}$ [7].

Details on the computation of these matrix elements can be found in the appendix A.

1.4 Impedance boundary conditions

In the section 1.2.7, the surface integral equations have been derived for a dielectric scatterer. It is not trivial that the EFIE (1.73a) and the MFIE (1.73b) in the exterior domain are actually linearly dependent, so a system that discretizes only the exterior domain cannot obtain the correct solution. Therefore, the interior domain EFIE (1.75a) and MFIE (1.75b) also need to be included in the discretization. Classical

formulations such as the Poggio-Miller-Chang-Harrington-Wu-Tsai (PMCHWT) [8–10] formulation or the Müller formulation [11] linearly combine the discretizations of the exterior (1.74) and interior (1.76) problems to obtain a solution.

The fact that the discretization of the interior domain is required for scattering or far field problems is not convenient since the interior fields are not the quantities of interest. A way to get rid of the interior problem is to impose a boundary condition on the tangential traces of the fields that is assumed to hold on the boundary of the scatterer Γ . One of the most used boundary conditions is the Perfect Electric Conductor (PEC) boundary condition that imposes the tangential trace of the electric field $\hat{\mathbf{n}} \times \mathbf{E}$ to vanish on the boundary Γ . Using the surface currents (1.58), the following relation holds for all \mathbf{r} on the boundary of a PEC

$$\mathbf{M}_s(\mathbf{r}) = \mathbf{E}(\mathbf{r}) \times \hat{\mathbf{n}}(\mathbf{r}) = \mathbf{0}. \quad (1.113)$$

The PEC is an ideal model as it implies that the material of the scatterer is lossless, impenetrable, has an infinite conductivity, and has a surface impedance that is zero. Therefore, it is widely used to model good conductors. But there are cases where a non-zero impedance or a finite conductivity are more appropriate assumptions, in addition to the fact that a scatterer may be non-penetrable and lossy, so that another model is required in these cases.

In particular, a model that suits well is the Impedance Boundary Condition (IBC) [12] that enforces a proportionality relation between the tangential traces of the electric field $\hat{\mathbf{n}} \times \mathbf{E}$ and magnetic field $\hat{\mathbf{n}} \times \mathbf{H}$, the proportionality factor being the surface impedance z^{imp} (in Ω), so that for all $\mathbf{r} \in \Gamma$

$$\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}(\mathbf{r}) = z^{\text{imp}}(\mathbf{r}) \hat{\mathbf{n}}(\mathbf{r}) \times (\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}(\mathbf{r})) \quad (1.114)$$

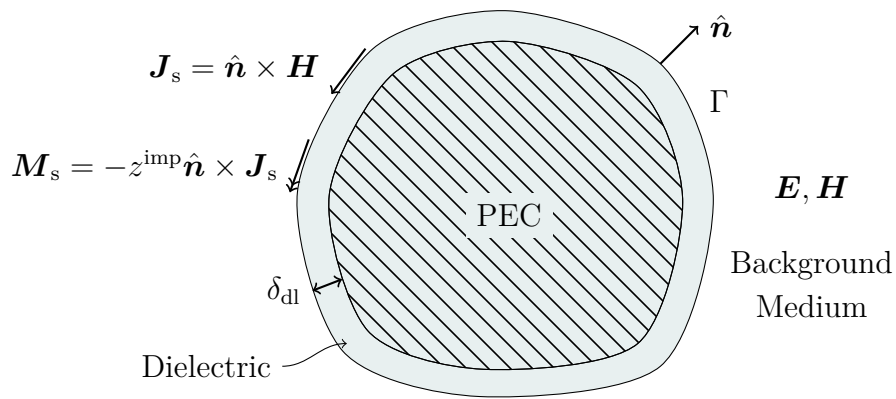
or equivalently, using the surface currents (1.58), the IBC reads

$$\mathbf{M}_s(\mathbf{r}) = -z^{\text{imp}}(\mathbf{r}) \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{J}_s(\mathbf{r}) \quad (1.115)$$

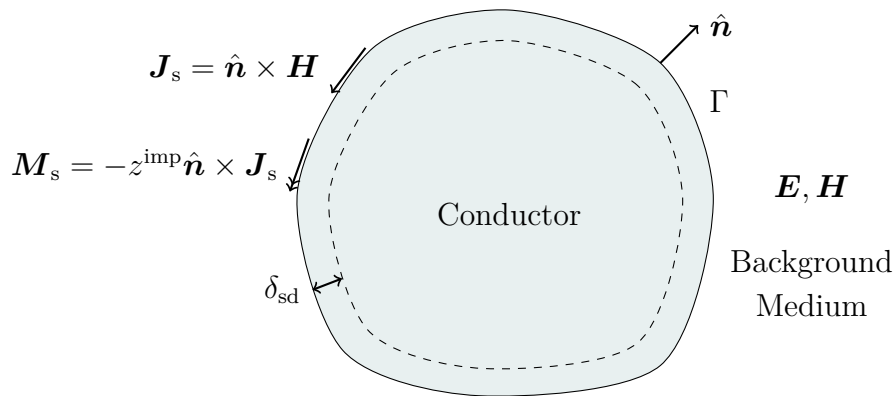
These IBCs (1.114) and (1.115) are usually referred as the Leontovich IBC [13]. The IBC can model a wide range of non-penetrable scatterers. In electromagnetics, there are two notable cases of such non-penetrable objects:

- A PEC coated by a dielectric layer (figure 1.10a): if the thickness of the dielectric layer δ_{dl} is small enough (compared to the radius of curvature of the scatterer), the currents flow in a thick layer between boundary of the scatterer (Γ) and the PEC.
- A conductor (figure 1.10b): in a conductor the physical current density decreases exponentially with the depth so that most of the current flows in a thick layer (the "skin") under the surface Γ which is the so-called skin effect. The characteristic length or skin depth δ_{sd} is defined as the depth where the current density is equal to $1/e \approx 36.8\%$ of the current density on the surface.

1. Background and notations



(a) PEC coated by a dielectric layer



(b) Skin effect on a conductor

Figure 1.10: Impedance Boundary Conditions on non-penetrable objects

In both cases, the currents flow in a thick layer right under the surface Γ . In the following, the skin depth δ is used indifferently to denote the thickness of the dielectric layer δ_{dl} or the characteristic length of the skin effect δ_{sd} .

The validity of the IBC has been vastly studied [14–18] and there are essentially two criteria to satisfy:

- The radius of curvature a of the scatterer should be large compared to the skin depth δ (i.e. $a \gg \delta$). Locally, a smooth surface can be considered as planar. This requirement on the smoothness of the surface of the scatterer is related to the fact that the derivations of the models for the impedance are usually done by studying the wave transmission when the geometry is an infinite plane. Therefore, the validity of the IBC on geometries that contain sharp corners or sharp edges is not guaranteed.
- The refraction index n' of the scatterer ($n' = c/v$ where c and v are the speeds of light in the background medium and in the scatterer) should be large enough (i.e. $n' \gg 1$) to consider that the transmitted waves are normal to the surface. It is related to the fact that the IBC does not take into account the incidence angle of the incident waves, so the derivations of the impedance models can be done with a normal incidence, and as a consequence the transmitted waves are also normal to the surface.

There are some generalizations of the IBC, most notably:

- Anisotropic impedance [19–21]. In this case the surface impedance is a tensor $\underline{\mathbf{Z}}^{\text{imp}}$, or equivalently there are two surface impedances z^{imp} and $z^{\text{imp}'}$ so that the IBC can be written as

$$\mathbf{M}_s(\mathbf{r}) = -\underline{\mathbf{Z}}^{\text{imp}}(\mathbf{r}) \cdot \mathbf{J}_s(\mathbf{r}) \quad (1.116a)$$

$$= -z^{\text{imp}}(\mathbf{r}) \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{J}_s(\mathbf{r}) - z^{\text{imp}' }(\mathbf{r}) \mathbf{J}_s(\mathbf{r}) \quad (1.116b)$$

Setting $z^{\text{imp}'} = 0$ results in the Leontovich IBC (1.115).

- Higher order IBC [22–24]. These IBCs add additional degrees of freedom in the characterization of the material by taking into account higher order terms such as the derivatives of the fields. Regarding to this, the Leontovich IBC (1.114) is the order 0, by using only the fields without higher derivatives.

It is not the concern of this work to determine whether a given model for the impedance is faithful or not in a given simulation. Instead, the surface impedance z^{imp} is used as an arbitrary input parameter in the formulations, the same way as the frequency, the permittivity, etc.

1.5 Preconditioning

The preconditioning of matrices stemming from the the MoM discretization of the surface integral operators derived in section 1.2.7 is an important part of this work. This section is split in three parts:

1. Background and notations

- The first part introduces the condition number and explains the importance of preconditioning in iterative solvers.
- The second part introduces the spectrum of the operators on a sphere. The PEC-EFIE is used as a reference example to illustrate the problems that arise from the MoM discretization.
- The third part introduces different tools that are used to solve the problems highlighted in the second part.

1.5.1 Condition number and preconditioning

1.5.1.1 Condition number

The condition number (CN) of an invertible matrix \mathbf{M} is defined as

$$\text{cond}(\mathbf{M}) = \|\mathbf{M}\| \cdot \|\mathbf{M}^{-1}\|. \quad (1.117)$$

The minimum condition number is 1. In general, the condition number depends on the norm chosen. In the following, the Euclidean norm $\|\cdot\|_2$ is assumed for $\|\cdot\|$ so that the norm of a matrix $\|\mathbf{M}\|$ is equal to its largest singular value $\sigma_{\max}(\mathbf{M})$. Similarly, the norm of its inverse $\|\mathbf{M}^{-1}\|$ is equal to the inverse of the minimum singular value $\sigma_{\min}(\mathbf{M})$. So, the condition number is equal to the ratio between the maximum and the minimum singular values of \mathbf{M}

$$\text{cond}(\mathbf{M}) = \frac{\sigma_{\max}(\mathbf{M})}{\sigma_{\min}(\mathbf{M})}. \quad (1.118)$$

These singular values can be obtained with a Singular Value Decomposition (SVD).

Assume that \mathbf{M} is the matrix representing a linear system to be solved

$$\mathbf{M}\mathbf{x} = \mathbf{y}. \quad (1.119)$$

If there is an error \mathbf{e} in the right hand side $\mathbf{y}_{\text{err}} = \mathbf{y} + \mathbf{e}$, the erroneous solution of the system is noted $\mathbf{x}_{\text{err}} = \mathbf{x} + \mathbf{M}^{-1}\mathbf{e}$ as

$$\mathbf{M}\mathbf{x}_{\text{err}} = \mathbf{y}_{\text{err}}. \quad (1.120)$$

The condition number is the maximum of the ratio between the relative error on \mathbf{x} and the relative error on \mathbf{y} :

$$\frac{\frac{\|\mathbf{x}_{\text{err}} - \mathbf{x}\|}{\|\mathbf{x}\|}}{\frac{\|\mathbf{y}_{\text{err}} - \mathbf{y}\|}{\|\mathbf{y}\|}} = \frac{\frac{\|\mathbf{M}^{-1}\mathbf{e}\|}{\|\mathbf{x}\|}}{\frac{\|\mathbf{e}\|}{\|\mathbf{M}\mathbf{x}\|}} = \frac{\|\mathbf{M}\mathbf{x}\|}{\|\mathbf{x}\|} \frac{\|\mathbf{M}^{-1}\mathbf{e}\|}{\|\mathbf{e}\|} \quad (1.121a)$$

$$\max \left(\frac{\|\mathbf{M}\mathbf{x}\|}{\|\mathbf{x}\|} \frac{\|\mathbf{M}^{-1}\mathbf{e}\|}{\|\mathbf{e}\|} \right) = \|\mathbf{M}\| \cdot \|\mathbf{M}^{-1}\| = \text{cond}(\mathbf{M}). \quad (1.121b)$$

In other words, the relative error on \mathbf{x} can be as high as the relative error on \mathbf{y} times the condition number. So, if the relative error on \mathbf{y} is in the order of the machine precision and the condition number is 10^N , it means that N digits of accuracy are lost on the solution \mathbf{x} . This loss of accuracy in the solution is independent of the method used to solve the system: the solver itself potentially introduces its own numerical issues and possibly won't converge at all if the condition number is too high. When the condition number of the system is too high (there is no sharp limit) the matrix is said ill-conditioned. It also implies that if the condition number of the matrix depends on a parameter (e.g. the frequency of the simulation) and is not bounded, then the matrix is also considered as ill-conditioned. On the contrary, if the condition number is low and stable when the simulation parameters vary, then the matrix is well-conditioned.

1.5.1.2 Preconditioning and iterative solvers

The goal of preconditioning is to find an invertible matrix \mathbf{L} such that

$$\text{cond}(\mathbf{LM}) < \text{cond}(\mathbf{M}) \quad (1.122)$$

in which case \mathbf{L} is a left preconditioner. Similarly, it is possible to look for an invertible matrix \mathbf{R} such that

$$\text{cond}(\mathbf{MR}) < \text{cond}(\mathbf{M}) \quad (1.123)$$

in which case \mathbf{R} is a right preconditioner. Additionally, it can be convenient to find both left and right preconditioners such that

$$\text{cond}(\mathbf{LMR}) < \text{cond}(\mathbf{M}). \quad (1.124)$$

For the left preconditioner case (1.122), the system solved is actually

$$\mathbf{LM}\mathbf{x} = \mathbf{L}\mathbf{y}. \quad (1.125)$$

For the right preconditioner case (1.123), the system solved is

$$\mathbf{MR}\mathbf{z} = \mathbf{y} \quad (1.126)$$

where the original unknown $\mathbf{x} = \mathbf{R}\mathbf{z}$ is retrieved when the auxiliary unknown \mathbf{z} has been found. Finally, using a left and right preconditioner (1.124)

$$\mathbf{LMR}\mathbf{z} = \mathbf{L}\mathbf{y} \quad (1.127)$$

is solved and as before $\mathbf{x} = \mathbf{R}\mathbf{z}$ is computed at the end when the system has been solved for \mathbf{z} .

Solving a linear system directly (e.g. by computing the inverse of the matrix) works when the number of unknowns N_s is low (e.g. $N_s \leq 10^4$, though it's not a sharp limit), but for much larger problems (e.g. $N_s \geq 10^5$), not only the classical cubic complexity ($\mathcal{O}(N_s^3)$) in time of the matrix inversion is prohibitive, but also the

1. Background and notations

quadratic complexity ($\mathcal{O}(N_s^2)$) in space needed for the storage of a dense matrix is a problem regarding the scalability of the solver.

In practice the solvers used are iterative. There are many existing iterative solvers, amongst the most known: GMRES (Generalized Minimal Residual), CG (Conjugate Gradient) and BiCG (Biconjugate Gradient). At each iteration n , an iterative solver computes an estimate of the solution $\tilde{\mathbf{x}}_n$ based on the previous estimate $\tilde{\mathbf{x}}_{n-1}$. The solver ends when the relative norm of the residual $\|\mathbf{M}\tilde{\mathbf{x}}_n - \mathbf{y}\|/\|\mathbf{y}\|$ is below a given threshold that depends on the accuracy wanted or when a certain number of iterations N_{iter} is reached. These solvers are all different in the way they work (some involve matrix-vector products with \mathbf{M} , some with $\overline{\mathbf{M}}^T$), on the required properties of \mathbf{M} (some requires it to be symmetric positive definite, others work with general non-symmetric matrices), on their stability or on the condition that ensure the convergence (in some specific cases the norm of the residual can be bounded by function exponentially decreasing with the number of iterations). GMRES is usually a good choice for non-symmetric matrices that have a low condition number. Note that the cost of GMRES increases at each iteration so it is usually restarted after a fixed number of iterations N_{res} with the last solution estimate $\tilde{\mathbf{x}}_{N_{\text{res}}}$ used as first guess $\tilde{\mathbf{x}}_0$ for the restarted solver.

In practice, for a large number of unknowns N_s , a time complexity $\mathcal{O}(N_s^2)$ for a matrix-vector product as well as a space complexity $\mathcal{O}(N_s^2)$ for the storage of the matrix is not considered as scalable. A fortiori, the direct matrix-matrix products are banned, so the preconditioned matrix **LMR** is never computed explicitly. Instead the product of **LMR** by a vector \mathbf{x} is computed from right to left as

$$\mathbf{LMR}\mathbf{x} = \mathbf{L}(\mathbf{M}(\mathbf{R}\mathbf{x})). \quad (1.128)$$

By noting N_{iter} the number of iterations, and by noting respectively $C_{\mathbf{L}\mathbf{v}}$, $C_{\mathbf{M}\mathbf{v}}$, $C_{\mathbf{R}\mathbf{v}}$ the complexity of a matrix-vector product with \mathbf{L} , \mathbf{M} , \mathbf{R} , the time complexity of solving the system (1.127) is $\mathcal{O}(N_{\text{iter}}(C_{\mathbf{L}\mathbf{v}} + C_{\mathbf{M}\mathbf{v}} + C_{\mathbf{R}\mathbf{v}}))$.

Fortunately, there are fast algorithms to compute matrix-vector products involving the MoM matrices defined in the section 1.3.4. The most notable algorithms are the Fast Multipole Method (FMM) [25] that reduces the space and time complexity to $C_{\mathbf{M}\mathbf{v}} = \mathcal{O}(N_s^{1.5})$ and the Multi-level FMM (MLFMM) [26, 27] that further reduces the complexity to $C_{\mathbf{M}\mathbf{v}} = \mathcal{O}(N_s \log(N_s))$.

Similarly, the asymptotic complexity of the preconditioner-vector multiplications as well as their memory storage complexity should not be more than the complexity of the original MoM (\mathbf{M}) matrix-vector multiplication and storage. In particular, as it is explained later, the Calderón preconditioning involves the discretization of the operator \mathcal{T} , in which case using the MLFMM results in $C_{\mathbf{L}\mathbf{v}} = \mathcal{O}(N_s \log(N_s))$. Similarly, for the right preconditioner, the aims should be a complexity such that $C_{\mathbf{R}\mathbf{v}} = \mathcal{O}(N_s \log(N_s))$.

Finally, regarding the number of iterations N_{iter} , the objective is to have a constant number of iterations for a given target accuracy so that the overall complexity of the solver is $\mathcal{O}(N_s \log(N_s))$. However, unless the preconditioned system has some specific properties, in general there is no explicit bound on the number of iterations. If a

relation between the condition number and N_{iter} is not available, the best is to keep the condition number stable and as low as possible, and demonstrate experimentally that the number of iterations is stable when the simulation input parameters are changed.

The main concern of this work is the reduction of the condition number of the matrices stemming from the MoM discretization in different formulations by using the appropriate preconditioners.

1.5.2 Analysis of the operators on a sphere

1.5.2.1 Spectrum of operators

In order to precondition the linear systems arising from the MoM discretization of the operators, it is important to understand the spectral properties of these operators. These operators can be studied only for few canonical geometries where an analytic expression is available.

In the following, the boundary of the scatterer Γ is a sphere of radius a . The spherical coordinate system (r, θ, φ) is used. To study a scalar field on the surface of a sphere, it is convenient to work with the Spherical Harmonics (SH)

$$Y_{lm}(\theta, \varphi) = Y_{lm}(\hat{\mathbf{r}}) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_{lm}(\cos\theta) e^{im\varphi} \quad (1.129)$$

where l and m are integers such that $-l \leq m \leq l$ and P_{lm} are the associated Legendre polynomials.

Similarly, the analysis of a vector field on a sphere is done with the Vector Spherical Harmonics (VSH) basis. The VSH are defined as

$$\mathbf{Y}_{lm}(\hat{\mathbf{r}}) = \hat{\mathbf{r}} Y_{lm}(\hat{\mathbf{r}}) \quad (1.130a)$$

$$\mathbf{X}_{lm}(\hat{\mathbf{r}}) = \frac{a}{i\sqrt{l(l+1)}} \hat{\mathbf{r}} \times \nabla Y_{lm}(\hat{\mathbf{r}}) \quad (1.130b)$$

$$\mathbf{U}_{lm}(\hat{\mathbf{r}}) = -\frac{a}{i\sqrt{l(l+1)}} \nabla Y_{lm}(\hat{\mathbf{r}}). \quad (1.130c)$$

The three VSH (1.130) form an orthonormal basis of the vectors fields on Γ with the inner product (1.81). As the surface integral operators \mathcal{T} (1.68a) and \mathcal{K} (1.69) work with the tangential traces of the fields on Γ , only the VSH \mathbf{X}_{lm} and \mathbf{U}_{lm} that are purely tangential are used. The VSH \mathbf{Y}_{lm} that is normal to Γ won't be used ($\hat{\mathbf{r}} = \hat{\mathbf{n}}$). Also, the order $l = 0$ is ignored as $\mathbf{X}_{00} = \mathbf{U}_{00} = \mathbf{0}$.

The \mathbf{X}_{lm} are solenoidal ($\nabla \cdot \mathbf{X}_{lm} = 0$), so the study of the behaviour of the operators applied on \mathbf{X}_{lm} infers their behaviours on the loop functions (1.96) and (1.102). Similarly, studying the operators applied on \mathbf{U}_{lm} infers their behaviours on the star functions (1.93).

The Riccati-Bessel and second kind Riccati-Hankel functions are noted J_l and $H_l^{(2)}$, and $'$ denotes a derivative. At a given wave number k , applying \mathcal{T} and \mathcal{K} to the

1. Background and notations

VSH results in the following analytic expression [28, 29]

$$\mathcal{T}\mathbf{X}_{lm} = -J_l(ka)H_l^{(2)}(ka)\mathbf{U}_{lm} \quad (1.131a)$$

$$\mathcal{T}\mathbf{U}_{lm} = J_l'(ka)H_l^{(2)'}(ka)\mathbf{X}_{lm} \quad (1.131b)$$

$$\mathcal{K}\mathbf{X}_{lm} = \frac{i}{2} \left(J_l(ka)H_l^{(2)'}(ka) + J_l'(ka)H_l^{(2)}(ka) \right) \mathbf{X}_{lm} \quad (1.131c)$$

$$\mathcal{K}\mathbf{U}_{lm} = -\frac{i}{2} \left(J_l(ka)H_l^{(2)'}(ka) + J_l'(ka)H_l^{(2)}(ka) \right) \mathbf{U}_{lm}. \quad (1.131d)$$

Additional useful operators are

$$\left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{U}_{lm} = iJ_l(ka)H_l^{(2)'}(ka)\mathbf{U}_{lm} \quad (1.132a)$$

$$\left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{X}_{lm} = -iJ_l'(ka)H_l^{(2)}(ka)\mathbf{X}_{lm} \quad (1.132b)$$

$$\left(\frac{\mathcal{I}}{2} + \mathcal{K} \right) \mathbf{U}_{lm} = -iJ_l'(ka)H_l^{(2)}(ka)\mathbf{U}_{lm} \quad (1.132c)$$

$$\left(\frac{\mathcal{I}}{2} + \mathcal{K} \right) \mathbf{X}_{lm} = iJ_l(ka)H_l^{(2)'}(ka)\mathbf{X}_{lm} \quad (1.132d)$$

$$\hat{\mathbf{n}} \times \mathbf{X}_{lm} = \mathbf{U}_{lm} \quad (1.132e)$$

$$\hat{\mathbf{n}} \times \mathbf{U}_{lm} = -\mathbf{X}_{lm}. \quad (1.132f)$$

Note that \mathbf{X}_{lm} and \mathbf{U}_{lm} are eigenfunctions of \mathcal{K} but not \mathcal{T} , in addition, the corresponding eigenvalues are not singular values either. In general, the singular values of a linear operator \mathcal{L} , are the square roots of the eigenvalues of the operator $\mathcal{L}^*\mathcal{L}$ where \mathcal{L}^* denotes the adjoint of \mathcal{L} . With some vector calculus manipulations, the following adjoints are obtained

$$\mathcal{T}^*\mathbf{f} = \hat{\mathbf{n}} \times \overline{\mathcal{T}(\hat{\mathbf{n}} \times \mathbf{f})} \quad (1.133a)$$

$$\mathcal{K}^*\mathbf{f} = \hat{\mathbf{n}} \times \overline{\mathcal{K}(\hat{\mathbf{n}} \times \mathbf{f})} \quad (1.133b)$$

$$(\hat{\mathbf{n}} \times \mathcal{I})^*\mathbf{f} = -\hat{\mathbf{n}} \times \mathbf{f}. \quad (1.133c)$$

Then, these adjoints are used to get the singular values of the operators. For example, for the operator \mathcal{T} (PEC-EFIE),

$$\mathcal{T}^*\mathcal{T}\mathbf{X}_{lm} = \left| J_l(ka)H_l^{(2)}(ka) \right|^2 \mathbf{X}_{lm} \quad (1.134a)$$

$$\mathcal{T}^*\mathcal{T}\mathbf{U}_{lm} = \left| J_l'(ka)H_l^{(2)'}(ka) \right|^2 \mathbf{U}_{lm}. \quad (1.134b)$$

So, the singular values of the operator \mathcal{T} are

$$\sigma_{\mathbf{X},l,m} = \left| J_l(ka)H_l^{(2)}(ka) \right| \quad (1.135a)$$

$$\sigma_{\mathbf{U},l,m} = \left| J_l'(ka)H_l^{(2)'}(ka) \right| \quad (1.135b)$$

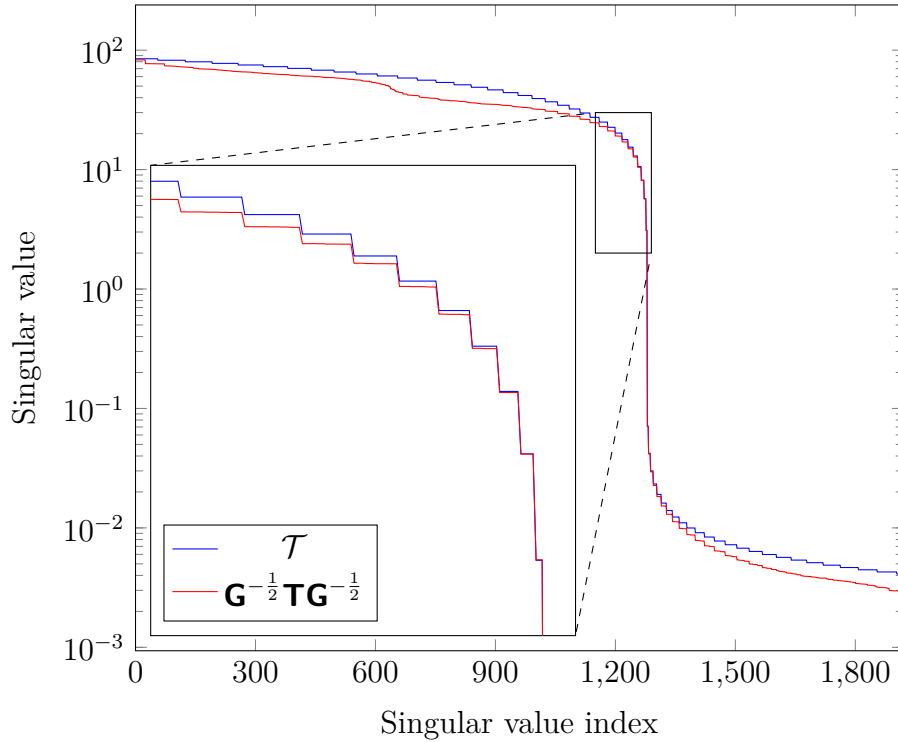


Figure 1.11: Singular values of the continuous operator \mathcal{T} and its MoM discretization matrix \mathbf{T}

for all $l \geq 1$ and $m \in [-l, l]$. As m does not appear in the expression of the singular values it actually means that the singular values for a given l have a $2l+1$ multiplicity.

The operator \mathcal{T} that has an infinite number of singular values is discretized by the MoM matrix $\mathbf{T} \in \mathbb{C}^{N_s \times N_s}$ that has a finite number of singular values. In other words, the discretization truncates the spectrum of \mathcal{T} to a finite number of singular values equals to N_s . The spherical harmonics become more oscillatory the higher is the order l , so increasing the mesh density adds harmonics.

The singular values of the MoM matrix \mathbf{T} for a unit sphere ($N_s = 1920$, $a = 1$ m, $f = 10$ MHz) have been computed (with an SVD) and plotted on the figure 1.11. In addition to that, the singular values of \mathcal{T} have been plotted as given before. To fit with the singular values of \mathcal{T} , the singular values of $\mathbf{G}^{-\frac{1}{2}} \mathbf{T} \mathbf{G}^{-\frac{1}{2}}$ (\mathbf{G} is the Gram matrix (1.111)) are actually plotted instead of \mathbf{T} in order to remove the scaling introduced by the testing and source basis functions. Also, the singular values of \mathcal{T} were selected such that $N_v - 1$ of them correspond to singular values associated to \mathbf{X}_{lm} and $N_f - 1$ of them correspond to \mathbf{U}_{lm} . It is related to the Euler's formula ($N_v + N_f = N_s + 2$ on a simply connected geometry) and the fact that \mathbf{X}_{lm} are solenoidal like the vertex based loop functions (1.96) and \mathbf{U}_{lm} are non-solenoidal like the face based star functions (1.93). It can be seen on the figure 1.11 that for lower l , the singular values match almost perfectly, but on higher l ($l \geq 5$), there are some discrepancies, because the discretization is not dense enough to discretize accurately the VSH at these orders. However, the asymptotic behaviour remains present.

There is no guarantee that the behaviour of the operators on a sphere is the same on a general geometry. However, it provides a valuable insight on the properties of the operators that is useful in understanding the problems of ill-conditioning.

1.5.2.2 Low frequency breakdown

The behaviour of the operators at low frequency is studied by looking at the asymptotic behaviour of the singular values when $k \rightarrow 0$. The Bessel family functions appearing in (1.135) have the following asymptotic behaviours [30]

$$J_l(ka) \underset{k \rightarrow 0}{\sim} \frac{(ka)^{l+1}}{(2l+1)!!} \quad (1.136a)$$

$$J'_l(ka) \underset{k \rightarrow 0}{\sim} \frac{(l+1)(ka)^l}{(2l+1)!!} \quad (1.136b)$$

$$H_l^{(2)}(ka) \underset{k \rightarrow 0}{\sim} \frac{i(2l-1)!!}{(ka)^l} \quad (1.136c)$$

$$H_l^{(2)'}(ka) \underset{k \rightarrow 0}{\sim} -\frac{il(2l-1)!!}{(ka)^{l+1}}. \quad (1.136d)$$

So, the asymptotic behaviour of the singular values of \mathcal{T} associated to \mathbf{X}_{lm} and \mathbf{U}_{lm} are

$$\sigma_{\mathbf{X},l,m} = \left| J_l(ka) H_l^{(2)}(ka) \right| \underset{k \rightarrow 0}{\sim} \frac{ka}{2l+1} \propto k \quad (1.137a)$$

$$\sigma_{\mathbf{U},l,m} = \left| J'_l(ka) H_l^{(2)'}(ka) \right| \underset{k \rightarrow 0}{\sim} \frac{(l+1)l}{(2l+1)ka} \propto \frac{1}{k}. \quad (1.137b)$$

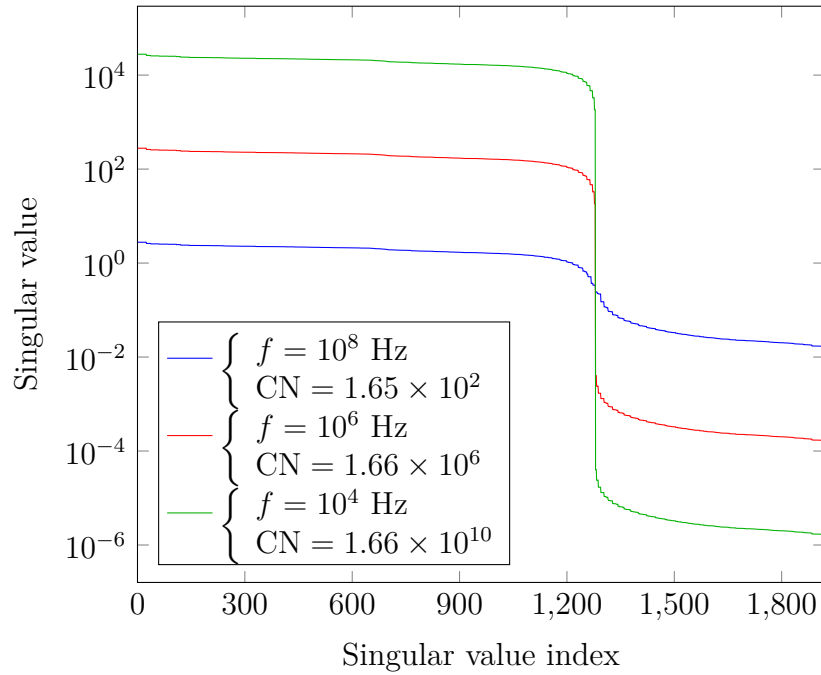
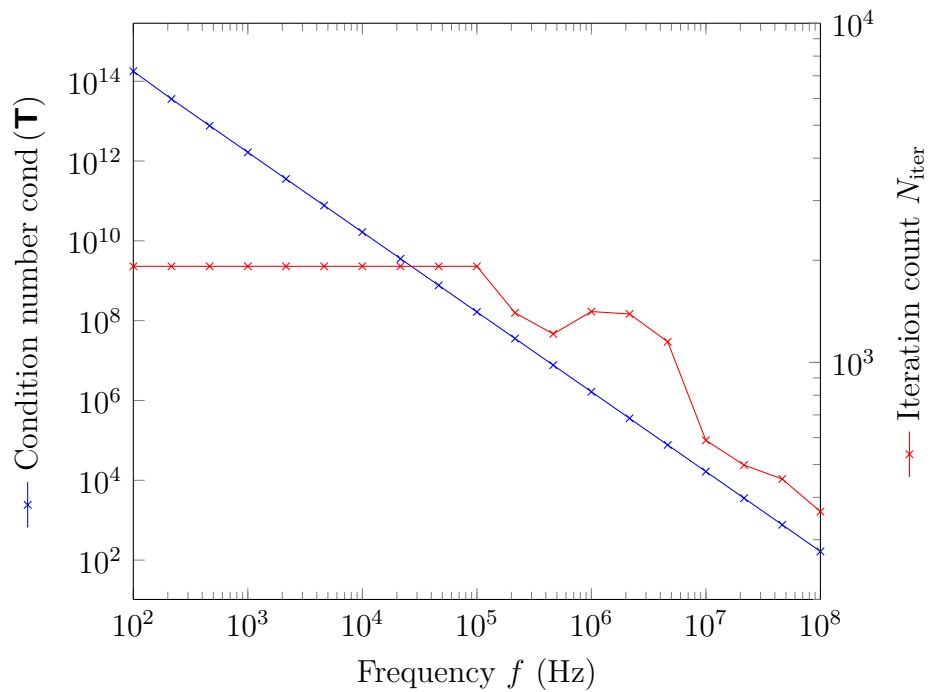
The proportionality symbol \propto is used asymptotically in the sense that when $k \rightarrow 0$, $f(k) \propto g(k) \iff f(k) = O(g(k))$ and $g(k) = O(f(k))$.

Since the singular values of \mathbf{T} have the same trend as (1.137) i.e. scale as k for solenoidal functions and $1/k$ for non-solenoidal functions, the condition number of \mathbf{T} scales proportionally to $1/k^2$ when $k \rightarrow 0$ i.e.

$$\text{cond}(\mathbf{T}) \underset{k \rightarrow 0}{\propto} \frac{1}{k^2}. \quad (1.138)$$

This is the so-called low frequency breakdown of the PEC-EFIE.

To illustrate the low frequency breakdown, the singular values of \mathbf{T} ($N_s = 1920$, $a = 1$ m) have been plotted for decreasing frequencies on the figure 1.12. It can be observed that the 2 branches effectively scale as $1/k$ and k . Dividing the frequency by 100 multiplies the condition number by 100^2 . In addition to that, the condition number has been plotted for several frequencies on the figure 1.13 where the quadratic growth of the condition number with the inverse of the frequency is clear. The number of iterations N_{iter} to reach a relative error of 10^{-8} on the residual with GMRES (restarted every 20 iterations) is also plotted. When there is a convergence $N_{\text{iter}} < N_s$, otherwise the solver has been stopped at $N_{\text{iter}} = N_s = 1920$ which indicates a non-convergence.

Figure 1.12: Singular values of \mathbf{T} for decreasing frequency f Figure 1.13: Condition number of \mathbf{T} and iteration count as functions of the frequency f

1. Background and notations

The low frequency breakdown is usually solved using a Calderón preconditioning (see section 1.5.3.1) or by separating the solenoidal and non-solenoidal contributions of the operator and rescale them independently to cancel the scaling in frequency (see section 1.5.3.3).

1.5.2.3 Dense mesh breakdown

As it has been explained in the section 1.5.2.1, increasing the mesh density (by increasing the number of edge N_s or decreasing the average edge length h) corresponds to adding harmonics (by increasing the order l).

As it has been done for the low frequency breakdown by letting $k \rightarrow 0$ in the previous section, the Bessel family functions appearing in (1.135) are studied for the dense mesh regime by letting the order $l \rightarrow +\infty$ (the following equivalents can be found using [30])

$$J_l(ka) \underset{l \rightarrow +\infty}{\sim} \frac{1}{\sqrt{2e}} \left(\frac{eka}{2l+1} \right)^{l+1} \quad (1.139a)$$

$$J'_l(ka) \underset{l \rightarrow +\infty}{\sim} \frac{l+1}{2l+1} \sqrt{\frac{e}{2}} \left(\frac{eka}{2l+1} \right)^l \quad (1.139b)$$

$$H_l^{(2)}(ka) \underset{l \rightarrow +\infty}{\sim} i \sqrt{\frac{2}{e}} \left(\frac{2l+1}{eka} \right)^l \quad (1.139c)$$

$$H_l^{(2)'}(ka) \underset{l \rightarrow +\infty}{\sim} -i \frac{l}{2l+1} \sqrt{2e} \left(\frac{2l+1}{eka} \right)^{l+1}. \quad (1.139d)$$

So, the asymptotic behaviour of the singular values of \mathcal{T} associated to \mathbf{X}_{lm} and \mathbf{U}_{lm} are

$$\sigma_{\mathbf{X},l,m} = \left| J_l(ka) H_l^{(2)}(ka) \right| \underset{l \rightarrow +\infty}{\sim} \frac{ka}{2l+1} \propto \frac{1}{l} \quad (1.140a)$$

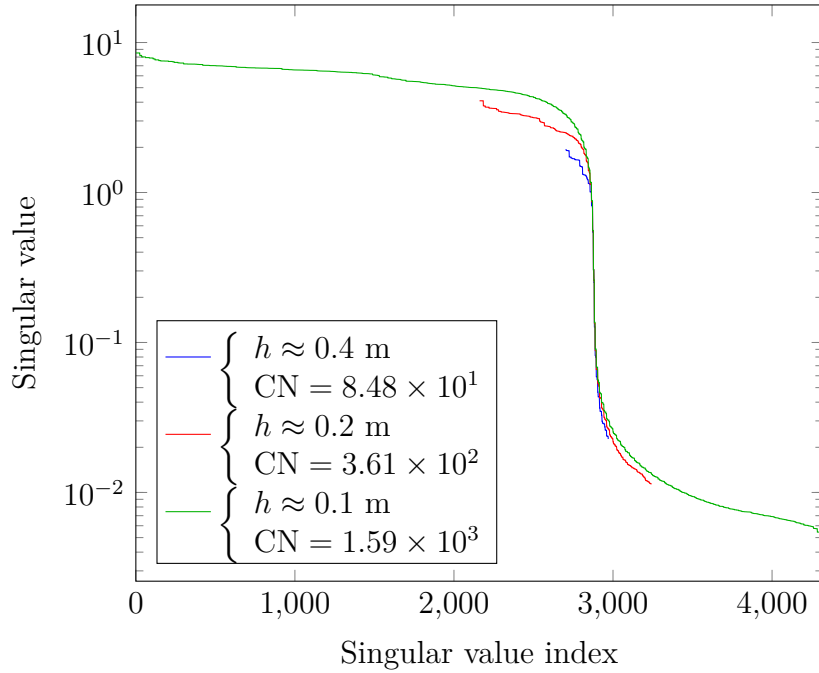
$$\sigma_{\mathbf{U},l,m} = \left| J'_l(ka) H_l^{(2)'}(ka) \right| \underset{l \rightarrow +\infty}{\sim} \frac{(l+1)l}{(2l+1)ka} \propto l. \quad (1.140b)$$

Assume that the discretization truncates the order of harmonics to some l_{\max} , so that there are $\sum_{l=1}^{l_{\max}} (2l+1) = l_{\max}^2 + 2l_{\max}$ harmonics. This number of harmonics corresponds to the number of edges N_s (up to some factor inferior to 1 as there are both \mathbf{X}_{lm} and \mathbf{U}_{lm}) so that

$$N_s \underset{h \rightarrow 0}{\propto} l_{\max}^2. \quad (1.141)$$

Consider that the sphere is discretized such that the average edge length is h and that the faces are triangles almost equilateral. Then, the number of faces (N_f) is equal to the area of the sphere ($\approx 4\pi a^2$) divided by the area of a face ($\approx h^2 \sqrt{3}/4$). So

$$N_f \underset{h \rightarrow 0}{\propto} \frac{a^2}{h^2}. \quad (1.142)$$

Figure 1.14: Singular values of \mathbf{T} for decreasing edge length h

Finally, as the mesh of the sphere is not open, there is exactly 2 faces for each edge, or equivalently if the 3 edges of each faces are summed then each edge is counted exactly twice, so

$$3N_f = 2N_s. \quad (1.143)$$

It results from (1.141), (1.142) and (1.143) that

$$l_{\max} \underset{h \rightarrow 0}{\propto} \frac{a}{h}. \quad (1.144)$$

Regarding the asymptotic behaviour (1.140) of the singular values of \mathcal{T} , the singular values of \mathbf{T} scale as h on the solenoidal functions and as $1/h$ on the non-solenoidal functions so

$$\text{cond}(\mathbf{T}) \underset{h \rightarrow 0}{\propto} \frac{1}{h^2}. \quad (1.145)$$

This is the so-called dense mesh (or dense grid) breakdown of the PEC-EFIE.

This dense mesh breakdown is illustrated on the figure 1.14 where the singular values of \mathbf{T} have been plotted for decreasing edge length. It can be observed that increasing the mesh density adds singular values on both branches scaling as $1/h$ and h . Dividing the edge length by 2 multiplies the condition number by roughly 4 ($= 2^2$). Again, the quadratic growth of the condition number with the inverse of the edge length is clear on the figure 1.15 where the condition number of \mathbf{T} has been plotted for decreasing edge length ($f = 100$ MHz, $a = 1$ m).

The dense mesh breakdown is usually solved using a preconditioning of type Calderón (see section 1.5.3.1), i.e. using a composition of operators to flatten the spectrum as opposed to a simple rescaling as it is done for the low frequency stabilization.

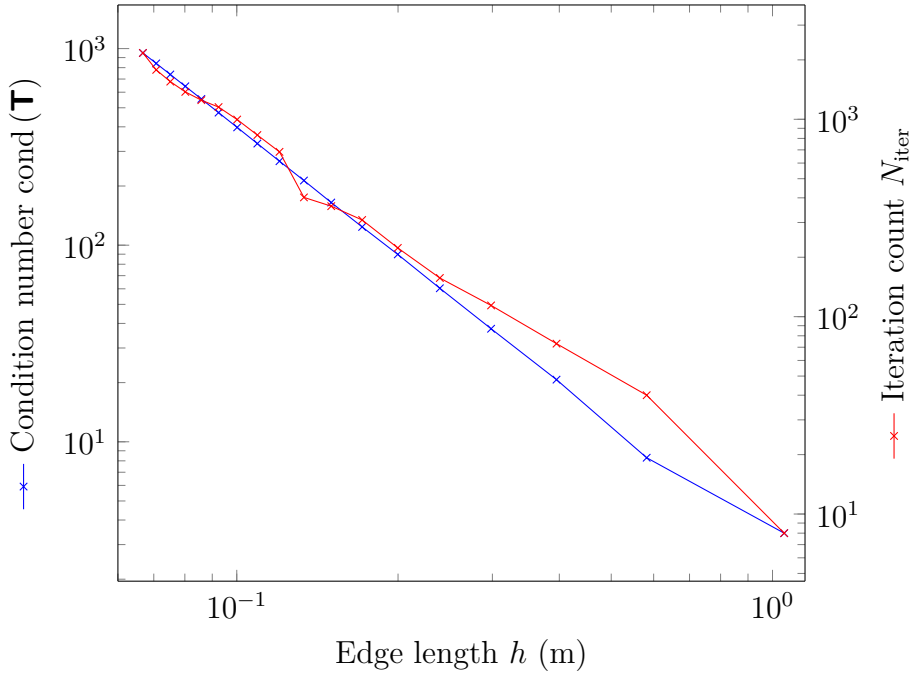


Figure 1.15: Condition number of \mathbf{T} and iteration count as functions of the average edge length h

1.5.2.4 Spurious resonances (high frequency)

At the opposite of low frequency (when $ka \ll 1$), there are also problems at high frequency (when $ka \gg 1$) but they are very different in their manifestations and in the way they are solved. Again, the Bessel family functions that appear in (1.135) are studied at high frequency by letting $k \rightarrow +\infty$

$$J_l(ka) = \sin\left(ka - l\frac{\pi}{2}\right) + O\left(\frac{1}{ka}\right) \quad (1.146a)$$

$$J'_l(ka) = \cos\left(ka - l\frac{\pi}{2}\right) + O\left(\frac{1}{ka}\right) \quad (1.146b)$$

$$H_l^{(2)}(ka) \underset{k \rightarrow +\infty}{\sim} i^{l+1} e^{-ika} \quad (1.146c)$$

$$H_l^{(2)'}(ka) \underset{k \rightarrow +\infty}{\sim} i^l e^{-ika}. \quad (1.146d)$$

It is clear from these asymptotic behaviours that the Riccati-Bessel functions J_l and their derivatives J'_l have zeros. These zeros are more frequent the higher is the frequency. It means that at some specific frequencies that correspond to the zeros of $J_l(ka)$ and $J'_l(ka)$, some singular values of \mathcal{T} are 0. As a result, around these frequencies the condition number of \mathbf{T} is unbounded.

On the figure 1.16, the condition number of \mathbf{T} ($N_s = 4320$, $a = 1$ m, $h \approx 0.1$ m) has been computed for the range of frequencies where the first resonances occur. The absolute value of the first orders of $J_l(ka)$ and $J'_l(ka)$ are superimposed to illustrate the correspondence of their zeros with the high condition numbers. It is clear that

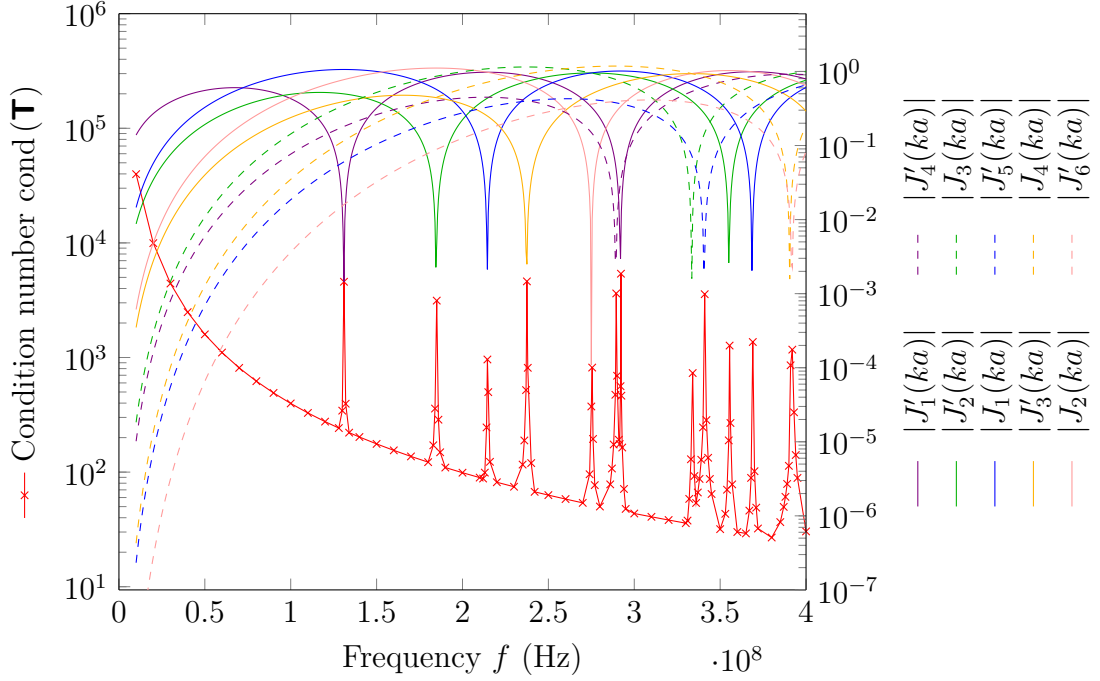


Figure 1.16: Condition number of \mathbf{T} as a function of the frequency f showing the first resonances that match with the zeros of the Riccati-Bessel functions and their derivatives

this behaviour worsens when the frequency increases as there are more and more zeros in the Riccati-Bessel functions and their derivatives.

These spurious resonances occur for the PEC-EFIE (\mathcal{T}) but they also occur for the PEC-MFIE ($\mathcal{I}/2 - \mathcal{K}$) at the same frequencies because the singular values of its operator are also proportional to $J_l(ka)$ and $J_l'(ka)$. In practice, this problem is solved by combining the EFIE and the MFIE, resulting in a CFIE (Combined Field Integral Equation). For example in the PEC case, the EFIE is

$$\eta \mathcal{T} \mathbf{J}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \quad (1.147)$$

and the MFIE is

$$\left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H}^{\text{inc}}. \quad (1.148)$$

Let α be a dimensionless factor. $(1 - \alpha)\hat{\mathbf{n}} \times$ is applied on the MFIE and summed with the EFIE scaled by α/η , such that the PEC-CFIE is

$$\left(\alpha \mathcal{T} + (1 - \alpha) \hat{\mathbf{n}} \times \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \right) \mathbf{J}_s = -\frac{\alpha}{\eta} \hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} + (1 - \alpha) \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{H}^{\text{inc}}). \quad (1.149)$$

1. Background and notations

Using (1.131) and (1.132), the VSH expansion of the PEC-CFIE operator is

$$\left(\alpha \mathcal{T} + (1 - \alpha) \hat{\mathbf{n}} \times \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \right) \mathbf{X}_{lm} = -(\alpha J_l(ka) + i(1 - \alpha) J_l'(ka)) H_l^{(2)}(ka) \mathbf{U}_{lm} \quad (1.150a)$$

$$\left(\alpha \mathcal{T} + (1 - \alpha) \hat{\mathbf{n}} \times \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \right) \mathbf{U}_{lm} = (\alpha J_l'(ka) - i(1 - \alpha) J_l(ka)) H_l^{(2)'}(ka) \mathbf{X}_{lm} \quad (1.150b)$$

and the singular values of the operator are

$$\sigma_{\mathbf{X},l,m} = |\alpha J_l(ka) + i(1 - \alpha) J_l'(ka)| |H_l^{(2)}(ka)| \quad (1.151a)$$

$$\sigma_{\mathbf{U},l,m} = |\alpha J_l'(ka) - i(1 - \alpha) J_l(ka)| |H_l^{(2)'}(ka)| \quad (1.151b)$$

for each $l \geq 1$, with a multiplicity $2l + 1$. If the factor α that combines the EFIE and the MFIE is chosen correctly ($\alpha = 1/2$ is a common choice), the singular values of the CFIE cannot be zero for any frequency which removes the spurious resonances. In this work however one of the main concerns is the low frequency, so these spurious resonances are not considered.

1.5.2.5 Numerical cancellations (low frequency)

Unfortunately, the ill-conditioning of the linear system is not the only problem that is present at low frequency (when $k \rightarrow 0$). Another problem is the numerical cancellation that occurs between the solenoidal and the non-solenoidal parts of the fields because they scale differently with the frequency.

Consider a solenoidal function \mathbf{f}_{sol} on Γ such that it exists a scalar function g that verifies $\mathbf{f}_{\text{sol}} = \hat{\mathbf{n}} \times \nabla g$ (e.g. $\boldsymbol{\lambda}_j$ (1.96) or \mathbf{X}_{lm} (1.130)). Then, because of $\partial\Gamma = \emptyset$,

$$\iint_{\mathbf{r} \in \Gamma} \mathbf{f}_{\text{sol}}(\mathbf{r}) \, dS = \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \nabla g(\mathbf{r}) \, dS = \oint_{\mathbf{r} \in \partial\Gamma} g(\mathbf{r}) \, d\mathbf{l} = \mathbf{0}. \quad (1.152)$$

In particular, for any constant field $\mathbf{C} \in \mathbb{R}^3$

$$\langle \mathbf{f}_{\text{sol}}, \mathbf{C} \rangle = 0, \quad (1.153)$$

where the inner product $\langle \cdot, \cdot \rangle$ is defined in (1.81). Note that some functions are solenoidal, i.e. $\nabla \cdot \mathbf{f}_{\text{sol}} = 0$, but do not verify $\mathbf{f}_{\text{sol}} = \hat{\mathbf{n}} \times \nabla g$, in particular the global loops $\boldsymbol{\eta}_j$ (1.102). In this case, (1.153) still holds true. It can be obtained by rewriting the constant field \mathbf{C} as the gradient of a scalar field and by using the duality between the gradient and the divergence on Γ

$$\langle \mathbf{f}_{\text{sol}}, \mathbf{C} \rangle = \iint_{\mathbf{r} \in \Gamma} \overline{\mathbf{f}_{\text{sol}}(\mathbf{r})} \cdot \mathbf{C} \, dS \quad (1.154a)$$

$$= \iint_{\mathbf{r} \in \Gamma} \overline{\mathbf{f}_{\text{sol}}(\mathbf{r})} \cdot \nabla(\mathbf{r} \cdot \mathbf{C}) \, dS \quad (1.154b)$$

$$= - \iint_{\mathbf{r} \in \Gamma} \overline{\nabla \cdot \mathbf{f}_{\text{sol}}(\mathbf{r})} (\mathbf{r} \cdot \mathbf{C}) \, dS \quad (1.154c)$$

$$= 0. \quad (1.154d)$$

Using this result, the asymptotic behaviour of the testing of a plane wave (1.42) by solenoidal functions is

$$\langle \mathbf{f}_{\text{sol}}, \mathbf{E}_{\text{pw}} \rangle = \langle \mathbf{f}_{\text{sol}}, E_0 e^{-i\mathbf{k} \cdot \mathbf{r}} \hat{\mathbf{p}} \rangle \quad (1.155a)$$

$$= E_0 \underbrace{\langle \mathbf{f}_{\text{sol}}, \hat{\mathbf{p}} \rangle}_{=0} - E_0 i k \underbrace{\langle \mathbf{f}_{\text{sol}}, (\hat{\mathbf{k}} \cdot \mathbf{r}) \hat{\mathbf{p}} \rangle}_{=O(k)} + E_0 \underbrace{\left\langle \mathbf{f}_{\text{sol}}, \hat{\mathbf{p}} \sum_{n=2}^{+\infty} \frac{(-i\mathbf{k} \cdot \mathbf{r})^n}{n!} \right\rangle}_{=O(k^2)} \quad (1.155b)$$

$$= O(k). \quad (1.155c)$$

In particular,

$$\langle \mathbf{X}_{lm}, \mathbf{E}_{\text{pw}} \rangle = O(k) \quad (1.156a)$$

$$\langle \mathbf{U}_{lm}, \hat{\mathbf{n}} \times \mathbf{E}_{\text{pw}} \rangle = O(k) \quad (1.156b)$$

$$\langle \mathbf{X}_{lm}, \hat{\mathbf{n}} \times \mathbf{E}_{\text{pw}} \rangle = O(1) \quad (1.156c)$$

$$\langle \mathbf{U}_{lm}, \mathbf{E}_{\text{pw}} \rangle = O(1). \quad (1.156d)$$

Therefore, in the testing of the RHS, when a plane wave is tested by a set of basis functions such as the RWG functions \mathbf{f}_n (1.90), each coefficient of the RHS vector contains the information for both the solenoidal and the non-solenoidal part of the field. When the frequency is lowered, at first the solenoidal part loses digits of accuracy until the frequency is low enough to get the solenoidal part completely cancelled by the non-solenoidal part. As the solenoidal and the non-solenoidal part contribute equally to the solution, the solution cannot be retrieved in such cases (independently of the conditioning of the system), unless a proper treatment is applied on the equation.

Similarly, it is common that the solenoidal and non-solenoidal parts of the currents scale differently. For example, it is shown in the section 1.5.3.2 that in the PEC-EFIE, with a plane wave excitation, the solenoidal part of the electric current scales as $O(1)$ and its non-solenoidal part scales as $O(k)$. Therefore, if they are summed together without precautions, the non-solenoidal part starts to lose digits of accuracy until it is completely cancelled at very low frequency. Again, this is a problem because the two parts contribute equally to the solution: in the far field the two contributions retrieve the same scaling. So the far field is completely wrong if a contribution has been cancelled.

The usual technique to avoid the cancellation is to separate the solenoidal and non-solenoidal components and rescale them independently with a loop-star/loop-tree decomposition (see section 1.5.3.2) or with the quasi-Helmholtz projectors (see section 1.5.3.3).

1.5.3 Solutions to the ill-conditioning

1.5.3.1 Calderón preconditioning

Consider the matrix form for the exterior problem (1.74) and rewrite it as

$$\begin{pmatrix} \frac{\mathcal{I}}{2} - \mathcal{K} & -\mathcal{T} \\ \mathcal{T} & \frac{\mathcal{I}}{2} - \mathcal{K} \end{pmatrix} \begin{pmatrix} \eta \mathbf{J}_s \\ \mathbf{M}_s \end{pmatrix} = \begin{pmatrix} \eta \hat{\mathbf{n}} \times \mathbf{H}^{\text{inc}} \\ -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \end{pmatrix}. \quad (1.157)$$

In the section 1.2.7, the surface Γ is the boundary of the scatterer. But in free space, the surface can be chosen anywhere and the medium is the same inside and outside ($\eta' = \eta$, $k' = k$). So in free space, the interior problem (1.76) can be rewritten similarly

$$\begin{pmatrix} \frac{\mathcal{I}}{2} + \mathcal{K} & \mathcal{T} \\ -\mathcal{T} & \frac{\mathcal{I}}{2} + \mathcal{K} \end{pmatrix} \begin{pmatrix} \eta \mathbf{J}_s \\ \mathbf{M}_s \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}. \quad (1.158)$$

These two operators in matrix form

$$\begin{pmatrix} \frac{\mathcal{I}}{2} - \mathcal{K} & -\mathcal{T} \\ \mathcal{T} & \frac{\mathcal{I}}{2} - \mathcal{K} \end{pmatrix} \text{ and } \begin{pmatrix} \frac{\mathcal{I}}{2} + \mathcal{K} & \mathcal{T} \\ -\mathcal{T} & \frac{\mathcal{I}}{2} + \mathcal{K} \end{pmatrix} \quad (1.159)$$

are the so-called Calderón projectors. They are complementary (their sum is the identity), so by using the properties of the projectors, their product is identically zero

$$\begin{pmatrix} \frac{\mathcal{I}}{2} - \mathcal{K} & -\mathcal{T} \\ \mathcal{T} & \frac{\mathcal{I}}{2} - \mathcal{K} \end{pmatrix} \begin{pmatrix} \frac{\mathcal{I}}{2} + \mathcal{K} & \mathcal{T} \\ -\mathcal{T} & \frac{\mathcal{I}}{2} + \mathcal{K} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (1.160)$$

The expansion of this product results in the Calderón identities

$$\mathcal{T}^2 = -\frac{\mathcal{I}}{4} + \mathcal{K}^2 \quad (1.161a)$$

$$\mathcal{T}\mathcal{K} = -\mathcal{K}\mathcal{T}. \quad (1.161b)$$

Note that these identities can be verified on a sphere using the VSH expansion of the operators (1.131) and (1.132). The first identity (1.161a) can be used to precondition efficiently \mathbf{T} because the operator \mathcal{T} applied on itself is a scaled identity ($-\mathcal{I}/4$) plus a compact operator (\mathcal{K}^2).

Another way to write \mathcal{T}^2 is simply by expanding it from its definition (1.68a)

$$\mathcal{T}^2 = -k^2 \mathcal{T}_s^2 - \mathcal{T}_s \mathcal{T}_h - \mathcal{T}_h \mathcal{T}_s \quad (1.162)$$

where

$$\mathcal{T}_h^2 = 0 \quad (1.163)$$

has been used. \mathcal{T}^2 is not available in closed form but it can be discretized by combining two conforming discretizations of \mathcal{T} as described in the section 1.3.2. A conforming discretization of \mathcal{T}^2 is [7]

$$\mathbb{T} \mathbf{G}_m^{-1} \mathbf{T} \quad (1.164)$$

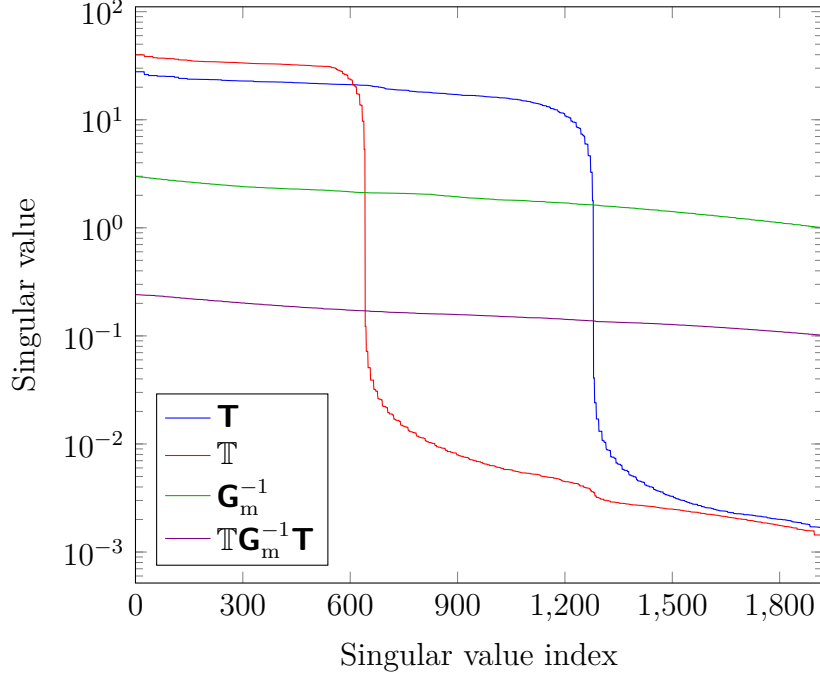


Figure 1.17: Singular values of $\mathbb{T}\mathbf{G}_m^{-1}\mathbf{T}$ that discretizes \mathcal{T}^2 ($\text{cond}(\mathbf{T}) = 1.65 \times 10^4$, $\text{cond}(\mathbb{T}) = 2.76 \times 10^4$, $\text{cond}(\mathbf{G}_m) = 2.98$, $\text{cond}(\mathbb{T}\mathbf{G}_m^{-1}\mathbf{T}) = 2.37$)

that is tested with rotated BC functions $\hat{\mathbf{n}} \times \mathbf{g}_n$ and that uses the RWG functions \mathbf{f}_n as sources. It would not be possible to use \mathbf{T} twice because the mix-Gram matrix whose elements are $\langle \hat{\mathbf{n}} \times \mathbf{f}_i, \mathbf{f}_j \rangle$ is not invertible.

In practice, the presence of the inverse of mix-Gram matrix \mathbf{G}_m^{-1} in the formulation is not a problem. Indeed, it is never computed explicitly: in an iterative solver only its product by a vector $\mathbf{y} = \mathbf{G}_m^{-1}\mathbf{x}$ is required which is actually computed by solving the system $\mathbf{G}_m\mathbf{y} = \mathbf{x}$ iteratively (so there is an iterative solver in the iterative solver). Because \mathbf{G}_m is sparse and well-conditioned for uniform discretizations, this system is solved iteratively in $\mathcal{O}(N_s)$ operations. For non-uniform discretizations, a diagonal preconditioning is used i.e. \mathbf{G}_m is preconditioned (on the left) by the diagonal matrix \mathbf{D} defined by

$$\mathbf{D}_{mn} = \begin{cases} \frac{1}{[\mathbf{G}_m]_{mn}} & \text{if } m = n \\ 0 & \text{otherwise.} \end{cases} \quad (1.165)$$

To illustrate the Calderón preconditioning, the spectra of \mathbb{T} , \mathbf{G}_m^{-1} , \mathbf{T} and their product $\mathbb{T}\mathbf{G}_m^{-1}\mathbf{T}$ have been plotted on the figure 1.17 for a unit sphere ($a = 1$ m, $N_s = 1920$, $f = 10$ MHz). It can be seen from the two branches of their spectra that \mathbb{T} and \mathbf{T} are both ill-conditioned, but the spectrum of $\mathbb{T}\mathbf{G}_m^{-1}\mathbf{T}$ is flat so this matrix is well-conditioned.

1.5.3.2 Loop-star decomposition

It is mentioned in the sections (1.5.2.2) and (1.5.2.5) that a way to solve the low frequency breakdown and the numerical cancellations is to separate the solenoidal and non-solenoidal contributions and rescale them independently. This separation can be done with a loop-star decomposition [5] using the Star-to-RWG basis $\mathbf{\Sigma} \in \mathbb{R}^{N_s \times N_f}$ (1.94) and the Loop-to-RWG basis $\mathbf{\Lambda} \in \mathbb{R}^{N_s \times N_v}$ (1.97) transformation matrices. According to the Euler's formula for planar graphs,

$$N_v + N_f = N_s + 2 \quad (1.166)$$

for simply connected geometries, so that all the N_v local loops and all the N_f stars basis functions cannot be linearly independent. For example, the all-one vectors $\mathbf{1}_{N_f} \in \mathbb{R}^{N_f}$ and $\mathbf{1}_{N_v} \in \mathbb{R}^{N_v}$ are respectively in the null spaces of $\mathbf{\Sigma}$ and $\mathbf{\Lambda}$

$$\mathbf{\Sigma} \mathbf{1}_{N_f} = \mathbf{0} \quad (1.167a)$$

$$\mathbf{\Lambda} \mathbf{1}_{N_v} = \mathbf{0}. \quad (1.167b)$$

In addition,

$$\mathbf{\Sigma}^T \mathbf{\Lambda} = \mathbf{0} \quad (1.168a)$$

$$\mathbf{\Lambda}^T \mathbf{\Sigma} = \mathbf{0}. \quad (1.168b)$$

To have a basis, 1 arbitrary local loop and 1 arbitrary star must be removed.

The multiply connected geometry are also considered in this work. Note N_h the number of handles (e.g. $N_h = 1$ on a torus). Then, the Euler's formula reads

$$N_v + N_f = N_s + 2 - 2N_h. \quad (1.169)$$

In that case, 1 local loop and 1 star basis must be removed and 2 global loops must be added for each handle. For example, these 2 global loops have been represented on a torus in the figure 1.9.

Let $\tilde{\mathbf{\Lambda}} \in \mathbb{R}^{N_s \times (N_v - N_h - 1)}$ and $\tilde{\mathbf{\Sigma}} \in \mathbb{R}^{N_s \times (N_f - N_h - 1)}$ be the local loop-to-RWG and star-to-RWG matrices where the adequate number of columns have been removed, and $\tilde{\mathbf{H}} \in \mathbb{R}^{N_s \times 2N_h}$ be the global loop-to-RWG matrix ($\tilde{\mathbf{H}}$ is used for consistency in the notations but it the same as \mathbf{H} in (1.102)). The loop-star decomposition matrix $\mathbf{M} \in \mathbb{R}^{N_s \times N_s}$ is then defined as the block matrix

$$\mathbf{M} = \begin{pmatrix} \tilde{\mathbf{\Lambda}} & \tilde{\mathbf{H}} & \tilde{\mathbf{\Sigma}} \end{pmatrix}. \quad (1.170)$$

Consider the PEC-EFIE

$$\eta \mathcal{T} \mathbf{J}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \quad (1.171)$$

discretized with RWG basis functions as

$$\eta \mathbf{T} \mathbf{J} = \mathbf{E} \quad (1.172)$$

where $\mathbf{T} = -ik\mathbf{T}_s + \frac{1}{ik}\mathbf{T}_h$ (see (1.103)) and

$$\mathbf{J}_s \approx \sum_{n=1}^{N_s} \mathbf{J}_n \mathbf{f}_n \quad (1.173a)$$

$$\mathbf{E}_m = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle. \quad (1.173b)$$

By introducing the coefficient vector \mathbf{Y} representing the current \mathbf{J} in a loop-star decomposed basis such that $\mathbf{J} = \mathbf{M}\mathbf{Y}$, (1.172) is equivalently rewritten in an loop-star decomposed form as

$$\eta \mathbf{M}^T \mathbf{T} \mathbf{M} \mathbf{Y} = \mathbf{M}^T \mathbf{E}. \quad (1.174)$$

Using the fact that the loop functions are solenoidal results in ($\mathbf{0}$ is the all-zero matrix but has a different dimension in each equation)

$$\mathbf{T}_h \tilde{\mathbf{\Lambda}} = \mathbf{0} \in \mathbb{R}^{N_s \times (N_v - N_h - 1)} \quad (1.175a)$$

$$\mathbf{T}_h \tilde{\mathbf{H}} = \mathbf{0} \in \mathbb{R}^{N_s \times 2N_h} \quad (1.175b)$$

$$\tilde{\mathbf{\Lambda}}^T \mathbf{T}_h = \mathbf{0} \in \mathbb{R}^{(N_v - N_h - 1) \times N_s} \quad (1.175c)$$

$$\tilde{\mathbf{H}}^T \mathbf{T}_h = \mathbf{0} \in \mathbb{R}^{2N_h \times N_s}. \quad (1.175d)$$

So, (1.174) is written in a block form as

$$\eta \begin{pmatrix} -ik\tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & -ik\tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{H}} & -ik\tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{\Sigma}} \\ -ik\tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & -ik\tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{H}} & -ik\tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{\Sigma}} \\ -ik\tilde{\mathbf{\Sigma}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & -ik\tilde{\mathbf{\Sigma}}^T \mathbf{T}_s \tilde{\mathbf{H}} & \tilde{\mathbf{\Sigma}}^T (-ik\mathbf{T}_s + \frac{1}{ik}\mathbf{T}_h) \tilde{\mathbf{\Sigma}} \end{pmatrix} \mathbf{Y} = \begin{pmatrix} \tilde{\mathbf{\Lambda}}^T \mathbf{E} \\ \tilde{\mathbf{H}}^T \mathbf{E} \\ \tilde{\mathbf{\Sigma}}^T \mathbf{E} \end{pmatrix}. \quad (1.176)$$

Using the asymptotic behaviour of the plane wave tested by solenoidal functions (1.155) results in the following scalings for the RHS

$$\mathbf{M}^T \mathbf{E} = \begin{pmatrix} \tilde{\mathbf{\Lambda}}^T \mathbf{E} \\ \tilde{\mathbf{H}}^T \mathbf{E} \\ \tilde{\mathbf{\Sigma}}^T \mathbf{E} \end{pmatrix} = \begin{pmatrix} \mathbf{O}(k) \\ \mathbf{O}(k) \\ \mathbf{O}(1) \end{pmatrix}. \quad (1.177)$$

Here the notation $\mathbf{O}(x)$ denotes a vector such that $\|\mathbf{O}(x)\| \underset{k \rightarrow 0}{\propto} x$. The sizes of the vectors is not indicated but it is clear that they are respectively the numbers of columns of $\tilde{\mathbf{\Lambda}}$, $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{\Sigma}}$. Similarly, in the following $\mathbf{O}(x)$ denotes a matrix such that $\|\mathbf{O}(x)\| \underset{k \rightarrow 0}{\propto} x$. The matrix on the LHS has the following asymptotic scaling of its blocks

$$\mathbf{M}^T \mathbf{T} \mathbf{M} = \begin{pmatrix} \mathbf{O}(k) & \mathbf{O}(k) & \mathbf{O}(k) \\ \mathbf{O}(k) & \mathbf{O}(k) & \mathbf{O}(k) \\ \mathbf{O}(k) & \mathbf{O}(k) & \mathbf{O}(k^{-1}) \end{pmatrix}. \quad (1.178)$$

This can be inverted, e.g. using the Schur complement, resulting in

$$(\mathbf{M}^T \mathbf{T} \mathbf{M})^{-1} = \begin{pmatrix} \mathbf{O}(k^{-1}) & \mathbf{O}(k^{-1}) & \mathbf{O}(k) \\ \mathbf{O}(k^{-1}) & \mathbf{O}(k^{-1}) & \mathbf{O}(k) \\ \mathbf{O}(k) & \mathbf{O}(k) & \mathbf{O}(k) \end{pmatrix}. \quad (1.179)$$

1. Background and notations

So, the scaling of the current in a loop-star decomposed basis is

$$\mathbf{Y} = \frac{1}{\eta} (\mathbf{M}^T \mathbf{T} \mathbf{M})^{-1} \mathbf{M}^T \mathbf{E} = \begin{pmatrix} \mathbf{O}(1) \\ \mathbf{O}(1) \\ \mathbf{O}(k) \end{pmatrix} \quad (1.180)$$

or equivalently,

$$\mathbf{J} = \mathbf{M} \mathbf{Y} = \underset{k \rightarrow 0}{\tilde{\mathbf{\Lambda}}} \mathbf{O}(1) + \underset{k \rightarrow 0}{\tilde{\mathbf{H}}} \mathbf{O}(1) + \underset{k \rightarrow 0}{\tilde{\mathbf{\Sigma}}} \mathbf{O}(k). \quad (1.181)$$

From this study, three problems can be noted:

- The solenoidal and non-solenoidal parts of the RHS \mathbf{E} scale differently which results in numerical cancellations (see section 1.5.2.5).
- Similarly, there are also numerical cancellations in the current \mathbf{J} .
- The matrix \mathbf{T} is ill-conditioned because one block scales as $\mathbf{O}(k^{-1})$ and the others scale as $\mathbf{O}(k)$, which results in a condition number that scales proportionally to k^{-2} .

By defining a left preconditioner in the form

$$\mathbf{L}^T = \begin{pmatrix} (-ika)^{-1} \tilde{\mathbf{\Lambda}}^T \\ (-ika)^{-1} \tilde{\mathbf{H}}^T \\ \tilde{\mathbf{\Sigma}}^T \end{pmatrix} \quad (1.182)$$

where a is in the order of the diameter of the scatterer, the numerical cancellation on the RHS is removed

$$\mathbf{L}^T \mathbf{E} = \begin{pmatrix} (-ika)^{-1} \tilde{\mathbf{\Lambda}}^T \mathbf{E} \\ (-ika)^{-1} \tilde{\mathbf{H}}^T \mathbf{E} \\ \tilde{\mathbf{\Sigma}}^T \mathbf{E} \end{pmatrix} = \begin{pmatrix} \mathbf{O}(1) \\ \mathbf{O}(1) \\ \mathbf{O}(1) \end{pmatrix}. \quad (1.183)$$

Then, with a right preconditioner in the form

$$\mathbf{R} = \begin{pmatrix} \tilde{\mathbf{\Lambda}} & \tilde{\mathbf{H}} & -ika \tilde{\mathbf{\Sigma}} \end{pmatrix} \quad (1.184)$$

the LHS matrix is

$$\eta \mathbf{L}^T \mathbf{T} \mathbf{R} = \eta \begin{pmatrix} \frac{1}{a} \tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & \frac{1}{a} \tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{H}} & -ik \tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{\Sigma}} \\ \frac{1}{a} \tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & \frac{1}{a} \tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{H}} & -ik \tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{\Sigma}} \\ -ik \tilde{\mathbf{\Sigma}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & -ik \tilde{\mathbf{\Sigma}}^T \mathbf{T}_s \tilde{\mathbf{H}} & \tilde{\mathbf{\Sigma}}^T (-ak^2 \mathbf{T}_s - a \mathbf{T}_h) \tilde{\mathbf{\Sigma}} \end{pmatrix} \quad (1.185a)$$

$$= \eta \begin{pmatrix} \frac{1}{a} \tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & \frac{1}{a} \tilde{\mathbf{\Lambda}}^T \mathbf{T}_s \tilde{\mathbf{H}} & \mathbf{0} \\ \frac{1}{a} \tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{\Lambda}} & \frac{1}{a} \tilde{\mathbf{H}}^T \mathbf{T}_s \tilde{\mathbf{H}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -a \tilde{\mathbf{\Sigma}}^T \mathbf{T}_h \tilde{\mathbf{\Sigma}} \end{pmatrix} + \underset{k \rightarrow 0}{\mathbf{O}}(k). \quad (1.185b)$$

The matrix has a limit in static whose non-zero blocks scale as $\mathbf{O}(1)$. In other words, the condition number of the system is stable and actually tends to the condition

number of the static limit matrix as $k \rightarrow 0$. So this preconditioning solves the low frequency breakdown. Then with this preconditioned system, the auxiliary current unknown does not have numerical cancellations

$$\mathbf{Y} = \frac{1}{\eta} (\mathbf{L}^\top \mathbf{T} \mathbf{R})^{-1} \mathbf{L}^\top \mathbf{E} = \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \\ \mathcal{O}(1) \end{pmatrix}. \quad (1.186)$$

To avoid the numerical cancellations on the RHS, the plane wave should be computed with the constant term removed manually on the solenoidal part i.e. the kernel $e^{-i\mathbf{k}\cdot\mathbf{r}}$ should be replaced by the extracted kernel $e^{-i\mathbf{k}\cdot\mathbf{r}} - 1$. Using an incident plane wave in the form $\mathbf{E}^{\text{inc}}(\mathbf{r}) = \mathbf{E}_0 e^{-i\mathbf{k}\cdot\mathbf{r}}$, the following equalities hold according to (1.155)

$$\tilde{\mathbf{A}}^\top \mathbf{E} = \tilde{\mathbf{A}}^\top \mathbf{E}_{\text{ext}} \quad (1.187a)$$

$$\tilde{\mathbf{H}}^\top \mathbf{E} = \tilde{\mathbf{H}}^\top \mathbf{E}_{\text{ext}} \quad (1.187b)$$

$$[\mathbf{E}]_m = -\langle \mathbf{f}_m, \mathbf{E}_0 e^{-i\mathbf{k}\cdot\mathbf{r}} \rangle \quad (1.187c)$$

$$[\mathbf{E}_{\text{ext}}]_m = -\langle \mathbf{f}_m, \mathbf{E}_0 (e^{-i\mathbf{k}\cdot\mathbf{r}} - 1) \rangle. \quad (1.187d)$$

However, only $\tilde{\mathbf{A}}^\top \mathbf{E}_{\text{ext}}$ and $\tilde{\mathbf{H}}^\top \mathbf{E}_{\text{ext}}$ provide the correct numerical results at very low frequency. $\tilde{\mathbf{A}}^\top \mathbf{E}$ and $\tilde{\mathbf{H}}^\top \mathbf{E}$ are plagued by a numerical cancellation.

Similarly, one should not attempt to compute $\mathbf{J} = \mathbf{R}\mathbf{Y}$ as it would reintroduce the cancellation. Instead, the solenoidal part $\mathbf{J}_{\mathbf{A}\mathbf{H}} = \tilde{\mathbf{A}}\mathcal{O}(1) + \tilde{\mathbf{H}}\mathcal{O}(1)$ and the non-solenoidal part $\mathbf{J}_{\mathbf{\Sigma}} = -ika\tilde{\mathbf{\Sigma}}\mathcal{O}(1)$ should remain separate. For example, the far field is computed by summing the field radiated by $\mathbf{J}_{\mathbf{A}\mathbf{H}}$ with the field radiated by $\mathbf{J}_{\mathbf{\Sigma}}$, instead of computing the field radiated by $\mathbf{J} = \mathbf{J}_{\mathbf{A}\mathbf{H}} + \mathbf{J}_{\mathbf{\Sigma}}$ that contains a cancellation of $\mathbf{J}_{\mathbf{\Sigma}}$. Similarly to the extraction performed on the RHS that enforces numerically the testing of a constant field by a solenoidal function to be 0 (see (1.153)), the same kind of extraction should be used for the scattering. In the case of the far field, the vector potential (1.60) is computed with an extracted kernel to radiate the solenoidal part of the current $\mathbf{J}_{\mathbf{A}\mathbf{H}}$

$$\mathbf{A}(\mathbf{r}) = \mu \frac{e^{-ik|\mathbf{r}|}}{4\pi|\mathbf{r}|} \iint_{\mathbf{r}' \in \Gamma} \left(e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}'} - 1 \right) \mathbf{J}_{\mathbf{A}\mathbf{H}}(\mathbf{r}') dS'. \quad (1.188)$$

Despite its advantages, there are two notable issues with the loop-star decomposition:

- It requires the detection of global loops which is computationally expensive in general. Therefore the use of a loop-star decomposition on multiply connected geometries is usually avoided.
- The loop-star decomposition enables the low-frequency stabilization and the removal of the numerical cancellations on the RHS and the current, but it introduces an ill-conditioning similar to the dense mesh breakdown. The ill-conditioning introduced by the loop-star basis results in a condition number

1. Background and notations

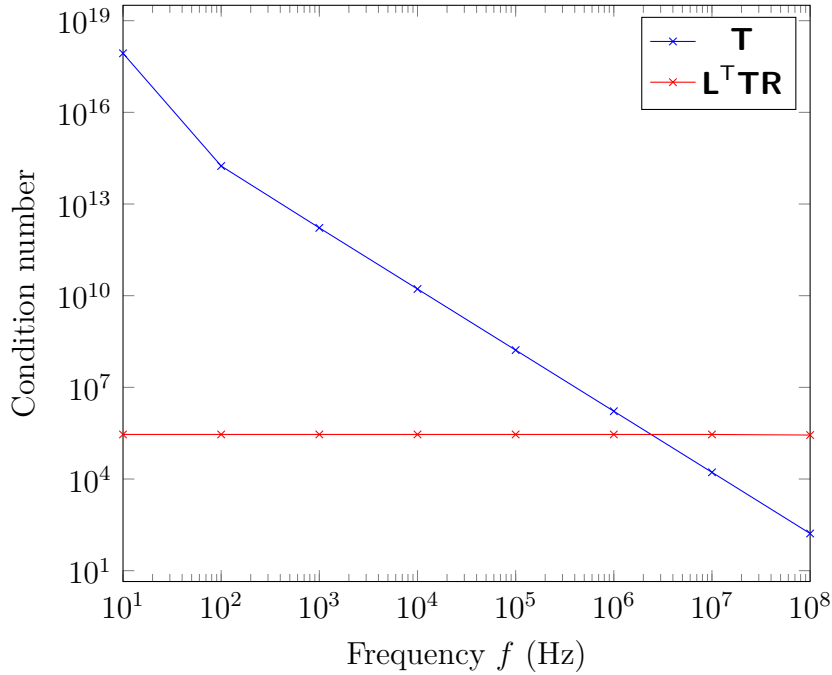


Figure 1.18: Condition number of \mathbf{T} and $\mathbf{L}^T \mathbf{TR}$ as functions of the frequency f

proportional to h^{-2} (h is the mesh edge length). This is on top of the dense mesh breakdown that also causes a condition number proportional to h^{-2} (see section 1.5.2.3), so in total it results in a condition number proportional to h^{-4} .

The figure 1.18 shows the condition numbers of \mathbf{T} and $\mathbf{L}^T \mathbf{TR}$ for decreasing frequencies ($h \approx 0.15$ m). It is clear that $\text{cond}(\mathbf{T}) \propto k^{-2}$ while $\text{cond}(\mathbf{L}^T \mathbf{TR})$ tends to a constant. In fact, this constant behaviour holds till arbitrary low frequency.

The figure 1.19 ($f = 10^6$ Hz) shows the degradation of the condition number in h introduced by the loop-star decomposition on which the scaling $\text{cond}(\mathbf{L}^T \mathbf{TR}) \propto h^{-4}$ is clear. So the system still suffers from the dense mesh breakdown that is actually worsened by the loop-star decomposition. At a fixed frequency, it can be seen that there is actually a point where $\text{cond}(\mathbf{L}^T \mathbf{TR}) > \text{cond}(\mathbf{T})$ if h is small enough.

Finally, the loop-star decomposition has been applied on vectors in a RWG basis, but it must be noted that the loop-star decomposition can also be applied on vectors in a BC basis. However, the loop-star decomposition has not the same interpretation when it is applied on a vector in a BC basis. Contrary to the decomposition in a RWG basis

$$\mathbf{J}^{\text{RWG}} = \tilde{\mathbf{\Lambda}} \mathbf{J}_{\mathbf{\Lambda}}^{\text{RWG}} + \tilde{\mathbf{H}} \mathbf{J}_{\mathbf{H}}^{\text{RWG}} + \tilde{\mathbf{\Sigma}} \mathbf{J}_{\mathbf{\Sigma}}^{\text{RWG}} \quad (1.189)$$

where $\tilde{\mathbf{\Lambda}} \mathbf{J}_{\mathbf{\Lambda}}^{\text{RWG}} + \tilde{\mathbf{H}} \mathbf{J}_{\mathbf{H}}^{\text{RWG}}$ corresponds to the solenoidal part and $\tilde{\mathbf{\Sigma}} \mathbf{J}_{\mathbf{\Sigma}}^{\text{RWG}}$ is non-solenoidal, in a BC basis

$$\mathbf{J}^{\text{BC}} = \tilde{\mathbf{\Lambda}} \mathbf{J}_{\mathbf{\Lambda}}^{\text{BC}} + \tilde{\mathbf{H}} \mathbf{J}_{\mathbf{H}}^{\text{BC}} + \tilde{\mathbf{\Sigma}} \mathbf{J}_{\mathbf{\Sigma}}^{\text{BC}} \quad (1.190)$$

the non-solenoidal part is $\tilde{\mathbf{\Lambda}} \mathbf{J}_{\mathbf{\Lambda}}^{\text{BC}}$ and the solenoidal part is $\tilde{\mathbf{H}} \mathbf{J}_{\mathbf{H}}^{\text{BC}} + \tilde{\mathbf{\Sigma}} \mathbf{J}_{\mathbf{\Sigma}}^{\text{BC}}$.

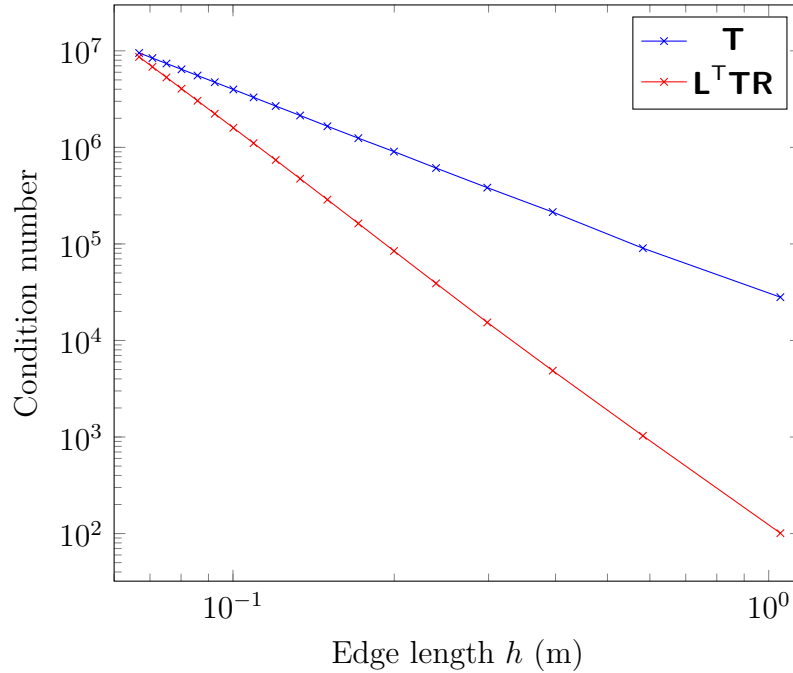


Figure 1.19: Condition number of \mathbf{T} and $\mathbf{L}^T \mathbf{TR}$ as functions of the average edge length h

1.5.3.3 Quasi-Helmholtz projectors

It is shown in the previous section that the loop-star decomposition enables a low frequency stabilization with some downsides. The quasi-Helmholtz projectors provide the same preconditioning benefits without the downsides. In particular, they enable the separation of the solenoidal and non-solenoidal parts without introducing a deterioration in the condition number and they do not require the detection of global loops.

Instead of using the matrix $\mathbf{\Sigma}$ directly, a projector $\mathbf{P}^\Sigma \in \mathbb{R}^{N_s \times N_s}$ orthogonal to its nullspace (the solenoidal functions) is built. Then, the projector $\mathbf{P}^{\Lambda H} \in \mathbb{R}^{N_s \times N_s}$ on the solenoidal subspace (that include both local and global loops) is its complementary

$$\mathbf{P}^\Sigma = \mathbf{\Sigma} (\mathbf{\Sigma}^\top \mathbf{\Sigma})^+ \mathbf{\Sigma}^\top \quad (1.191a)$$

$$\mathbf{P}^{\Lambda H} = \mathbf{I} - \mathbf{P}^\Sigma \quad (1.191b)$$

where $^+$ denotes the Moore-Penrose pseudo inverse.

As the projectors are dense, they are not computed or stored explicitly in practice. Instead, in iterative solvers their products with vectors are actually performed. By leveraging an algebraic multigrid, these products are computed in linear time as explained in [31].

As before, the PEC-EFIE (1.172)

$$\eta \mathbf{T} \mathbf{J} = \mathbf{E} \quad (1.192)$$

1. Background and notations

is used as a reference. Similarly to the loop-star decomposition matrix \mathbf{M} (1.170), the projectors can be used to visualize the different blocks of the operators and their scalings when the solenoidal (ΛH) or non-solenoidal (Σ) functions are used as source or testing. In the case of projectors, it is simply done by multiplying the identity $\mathbf{I} = \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma}$ on the left and the right of $\mathbf{T} = -ik\mathbf{T}_s + \frac{1}{ik}\mathbf{T}_h$ (1.103) and rewriting the resulting sum as a block matrix-vector products

$$\mathbf{T} = (\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma}) \mathbf{T} (\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma}) \quad (1.193a)$$

$$= -ik\mathbf{P}^{\Lambda H}\mathbf{T}_s\mathbf{P}^{\Lambda H} - ik\mathbf{P}^{\Lambda H}\mathbf{T}_s\mathbf{P}^{\Sigma} - ik\mathbf{P}^{\Sigma}\mathbf{T}_s\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \left(-ik\mathbf{T}_s + \frac{1}{ik}\mathbf{T}_h \right) \mathbf{P}^{\Sigma} \quad (1.193b)$$

$$= \begin{pmatrix} \mathbf{P}^{\Lambda H} & \mathbf{P}^{\Sigma} \end{pmatrix} \begin{pmatrix} -ik\mathbf{T}_s & -ik\mathbf{T}_s \\ -ik\mathbf{T}_s & -ik\mathbf{T}_s + \frac{1}{ik}\mathbf{T}_h \end{pmatrix} \begin{pmatrix} \mathbf{P}^{\Lambda H} \\ \mathbf{P}^{\Sigma} \end{pmatrix} \quad (1.193c)$$

where similarly to (1.175), the following equalities have been used

$$\mathbf{T}_h\mathbf{P}^{\Lambda H} = \mathbf{P}^{\Lambda H}\mathbf{T}_h = \mathbf{0}. \quad (1.194)$$

Note that the following equality holds as a consequence

$$\mathbf{P}^{\Sigma}\mathbf{T}_h\mathbf{P}^{\Sigma} = \mathbf{T}_h. \quad (1.195)$$

The figure 1.20 shows the effects of the projectors on the spectrum of \mathbf{T} ($f = 10^8$ Hz, $h \approx 0.15$ m). It can be seen that the projectors are used to select a branch. The preconditioners can take advantage of this to rescale a branch independently from the other. Then, using an approach similar to the preconditioning done with the loop-star decomposition in the section 1.5.3.2, the following left and right preconditioner (that are analogous to (1.182) and (1.184)) are constructed to precondition the PEC-EFIE

$$\mathbf{L} = -\frac{1}{ika}\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \quad (1.196a)$$

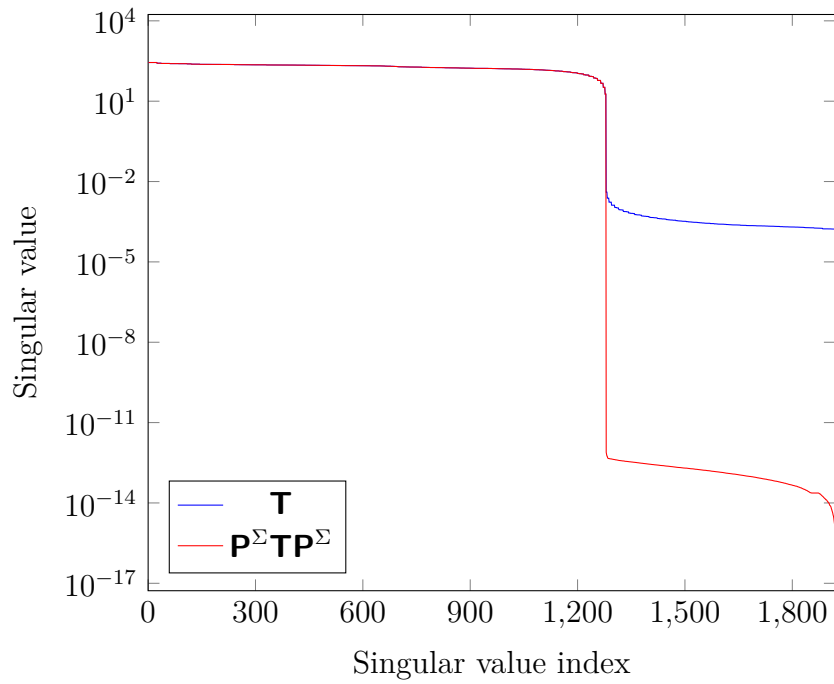
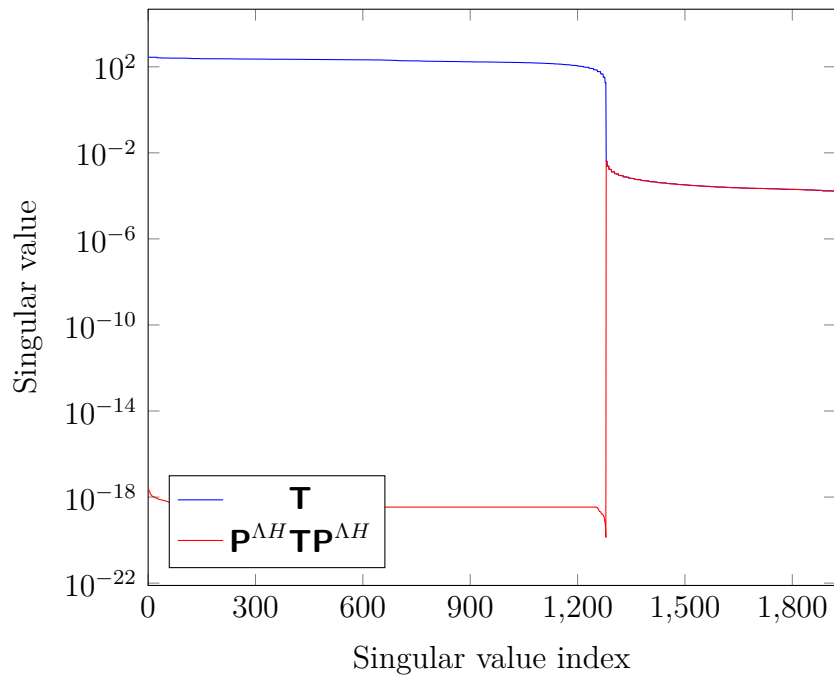
$$\mathbf{R} = \mathbf{P}^{\Lambda H} - ik a \mathbf{P}^{\Sigma}. \quad (1.196b)$$

Thus, the low frequency stable PEC-EFIE that is free from numerical cancellations on the RHS (plane wave) and the current is

$$\eta\mathbf{L}\mathbf{T}\mathbf{R}\mathbf{Y} = \mathbf{L}\mathbf{E} \quad (1.197a)$$

$$\mathbf{R}\mathbf{Y} = \mathbf{J}. \quad (1.197b)$$

Again, the preconditioned system is written as a block matrix-vector product to

(a) Singular values of \mathbf{T} and $\mathbf{P}^\Sigma \mathbf{T} \mathbf{P}^\Sigma$ (b) Singular values of \mathbf{T} and $\mathbf{P}^{\Lambda H} \mathbf{T} \mathbf{P}^{\Lambda H}$ Figure 1.20: Effect of the quasi-Helmholtz projectors on the spectrum of \mathbf{T}

1. Background and notations

highlight the different contributions

$$\mathbf{LTR} = \left(-\frac{1}{ika} \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \right) \mathbf{T} (\mathbf{P}^{\Lambda H} - ika \mathbf{P}^{\Sigma}) \quad (1.198a)$$

$$= \frac{1}{a} \mathbf{P}^{\Lambda H} \mathbf{T}_s \mathbf{P}^{\Lambda H} - ik \mathbf{P}^{\Lambda H} \mathbf{T}_s \mathbf{P}^{\Sigma} - ik \mathbf{P}^{\Sigma} \mathbf{T}_s \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} (-k^2 a \mathbf{T}_s - a \mathbf{T}_h) \mathbf{P}^{\Sigma} \quad (1.198b)$$

$$= \frac{1}{a} \mathbf{P}^{\Lambda H} \mathbf{T}_s \mathbf{P}^{\Lambda H} - a \mathbf{T}_h + \mathbf{O}_{k \rightarrow 0}(k) \quad (1.198c)$$

$$= \begin{pmatrix} \mathbf{P}^{\Lambda H} & \mathbf{P}^{\Sigma} \end{pmatrix} \begin{pmatrix} \frac{1}{a} \mathbf{T}_s & -ik \mathbf{T}_s \\ -ik \mathbf{T}_s & -k^2 a \mathbf{T}_s - a \mathbf{T}_h \end{pmatrix} \begin{pmatrix} \mathbf{P}^{\Lambda H} \\ \mathbf{P}^{\Sigma} \end{pmatrix} \quad (1.198d)$$

$$= \begin{pmatrix} \mathbf{P}^{\Lambda H} & \mathbf{P}^{\Sigma} \end{pmatrix} \begin{pmatrix} \frac{1}{a} \mathbf{T}_s & \mathbf{0} \\ \mathbf{0} & -a \mathbf{T}_h \end{pmatrix} \begin{pmatrix} \mathbf{P}^{\Lambda H} \\ \mathbf{P}^{\Sigma} \end{pmatrix} + \mathbf{O}_{k \rightarrow 0}(k). \quad (1.198e)$$

However, the system is still suffering from the dense mesh breakdown. The idea to solve this breakdown is to use a Calderón-like preconditioning in the sense that operators spectral properties are used to precondition the system. As it can be observed on the figure 1.14, when the edge length h of the mesh decreases, the dense mesh breakdown is the result of a growing of the maximum singular value proportional to h^{-1} on the non-solenoidal part of the spectrum (dominated by \mathcal{T}_h) and a diminution of the minimum singular proportional to h on the solenoidal part of the spectrum (only \mathcal{T}_s). In fact, applying \mathcal{T}_h on \mathcal{T}_s , and conversely applying \mathcal{T}_s on \mathcal{T}_h , removes the ill-conditioning in h . Regarding the Calderón identity (1.161a) and the expansion of \mathcal{T}^2 (1.162), it is in some measure what takes place in the classical Calderón preconditioning.

Similarly to the procedure done in the section 1.5.3.1 for the classical Calderón preconditioning, the inverse of the mix-Gram matrix is used to combine the primary system discretized with RWG basis functions to the preconditioner discretized with BC basis functions. The action of the quasi-Helmholtz projectors on a vector in the BC basis is different than their action on a vector in the RWG basis. In fact, regarding the loop-star decompositions (1.189) and (1.190), the columns of $\mathbf{\Sigma}$ are non-solenoidal and the columns of $\mathbf{\Lambda}$ and \mathbf{H} are solenoidal in RWG basis, as opposed to the BC basis in which the columns of $\mathbf{\Sigma}$ and \mathbf{H} are solenoidal and the columns of $\mathbf{\Lambda}$ are non-solenoidal. This suggest the introduction of the dual projectors \mathbf{P}^{Λ} and $\mathbf{P}^{\Sigma H}$ that are constructed the same way as \mathbf{P}^{Σ} and $\mathbf{P}^{\Lambda H}$. The dual projectors are

$$\mathbf{P}^{\Lambda} = \mathbf{\Lambda} (\mathbf{\Lambda}^{\top} \mathbf{\Lambda})^+ \mathbf{\Lambda}^{\top} \quad (1.199a)$$

$$\mathbf{P}^{\Sigma H} = \mathbf{I} - \mathbf{P}^{\Lambda}. \quad (1.199b)$$

Note that on simply connected geometry there are no global loops so $\mathbf{P}^{\Sigma} = \mathbf{P}^{\Sigma H}$ and $\mathbf{P}^{\Lambda} = \mathbf{P}^{\Lambda H}$. Useful identities coming from the properties of \mathcal{T}_h that are analogous to (1.194) and (1.195) are

$$\mathbf{T}_h \mathbf{P}^{\Sigma H} = \mathbf{P}^{\Sigma H} \mathbf{T}_h = \mathbf{0} \quad (1.200a)$$

$$\mathbf{P}^{\Lambda} \mathbf{T}_h \mathbf{P}^{\Lambda} = \mathbf{T}_h. \quad (1.200b)$$

In addition, the properties of the mix-Gram matrix \mathbf{G}_m imply the following identities that are analogous to (1.168b) and (1.163)

$$\mathbf{P}^\Lambda \mathbf{G}_m^{-1} \mathbf{P}^\Sigma = \mathbf{0} \quad (1.201a)$$

$$\mathbb{T}_h \mathbf{G}_m^{-1} \mathbf{T}_h = \mathbf{0}. \quad (1.201b)$$

The dense mesh preconditioner $\tilde{\mathbb{T}}$ is

$$\tilde{\mathbb{T}} = \frac{1}{a} \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{P}^{\Sigma H} - a \mathbb{T}_h. \quad (1.202)$$

This preconditioner can be seen as the low frequency approximation of

$$\left(-\frac{1}{ika} \mathbf{P}^{\Sigma H} + \mathbf{P}^\Lambda \right) \mathbb{T} (\mathbf{P}^{\Sigma H} - ika \mathbf{P}^\Lambda) = \tilde{\mathbb{T}} + \mathbf{O}_{k \rightarrow 0}(k), \quad (1.203)$$

so it is effectively achieving a Calderón preconditioning at arbitrary low frequency. The low frequency and dense mesh stable PEC-EFIE that is free of numerical cancellations reads

$$\eta \tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{T} \mathbf{R} \mathbf{Y} = \tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{E} \quad (1.204a)$$

$$\mathbf{R} \mathbf{Y} = \mathbf{J}. \quad (1.204b)$$

This formulation is capable of computing the solution current for an arbitrary low frequency and an arbitrary dense mesh. But in practice, some precautions must be taken to avoid numerical cancellations. The matrix products that yield $\mathbf{0}$ in (1.200) and (1.201) must be enforced by expanding the matrix products in (1.204) and actually removing the zero terms from the computation. The LHS matrix is computed as

$$\tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{T} \mathbf{R} = -\frac{ik}{a} \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{P}^{\Sigma H} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{T}_s \mathbf{R} \quad (1.205a)$$

$$- \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{G}_m^{-1} \mathbf{T}_h \quad (1.205b)$$

$$- \mathbb{T}_h \mathbf{G}_m^{-1} \mathbf{T}_s \mathbf{R}. \quad (1.205c)$$

The RHS is computed using the extracted kernel \mathbf{E}_{ext} as in the loop-star decomposition (1.187). With the projectors, $\mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}} = \mathbf{P}^{\Lambda H} \mathbf{E}$ holds true, but at low frequency only $\mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}}$ provides the correct result because $\mathbf{P}^{\Lambda H} \mathbf{E}$ has cancellations. The RHS is computed as

$$\tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{E} = \frac{1}{a} \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{G}_m^{-1} \mathbf{P}^\Sigma \mathbf{E} \quad (1.206a)$$

$$- \frac{1}{ika^2} \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{P}^{\Sigma H} \mathbf{G}_m^{-1} \mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}} \quad (1.206b)$$

$$+ \frac{1}{ik} \mathbb{T}_h \mathbf{G}_m^{-1} \mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}}. \quad (1.206c)$$

1. Background and notations

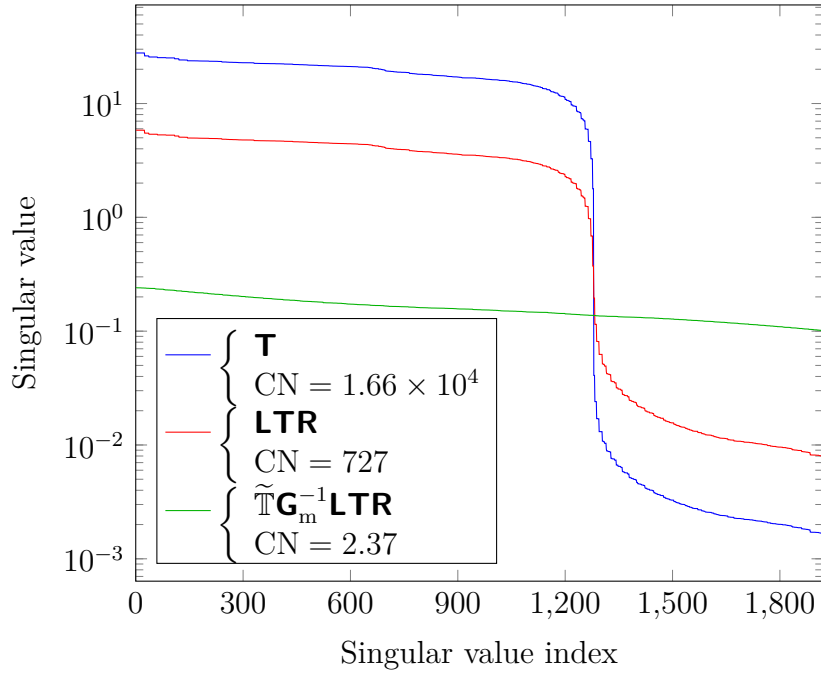


Figure 1.21: Singular values of the non-preconditioned \mathbf{T} , the low frequency preconditioned \mathbf{LTR} , and the fully preconditioned $\tilde{\mathbf{T}}\mathbf{G}_m^{-1}\mathbf{LTR}$

Finally, the solution is retrieved as

$$\mathbf{J} = \mathbf{R}\mathbf{Y} = \mathbf{J}_{\Lambda\mathbf{H}} + \mathbf{J}_{\Sigma} \quad (1.207a)$$

$$\mathbf{J}_{\Lambda\mathbf{H}} = \mathbf{P}^{\Lambda H}\mathbf{Y} \quad (1.207b)$$

$$\mathbf{J}_{\Sigma} = -ika\mathbf{P}^{\Sigma}\mathbf{Y}. \quad (1.207c)$$

Again, if the quantity of interest is the scattered field, then it should be computed as the sum of the fields scattered by $\mathbf{J}_{\Lambda\mathbf{H}}$ and \mathbf{J}_{Σ} rather than the field scattered by $\mathbf{J}_{\Lambda\mathbf{H}} + \mathbf{J}_{\Sigma}$ to avoid numerical cancellations. Similarly to the loop-star decomposition, the solenoidal part of the current $\mathbf{J}_{\Lambda\mathbf{H}}$ should be radiated using the extracted kernel in the vector potential (1.188).

The figure 1.21 (unit sphere, $f = 10^7$ Hz, $h \approx 0.15$ m) shows the effects of the low frequency preconditioner and the dense mesh preconditioner on the spectrum of the system. Roughly speaking, the effect of the low frequency preconditioner is to bring closer the two branches with a rescaling in frequency while the effect of the dense mesh preconditioner is to flatten the spectrum with a Calderón preconditioning.

The figure 1.22 (unit sphere, $f = 10^{-20}$ Hz, $h \approx 0.15$ m, $\mathbf{E}^{\text{inc}}(\mathbf{r}) = e^{ik\hat{z}\cdot\mathbf{r}}\hat{\mathbf{x}}$ V/m) shows the RCS ($\varphi = 0^\circ$, $\theta \in [0, 180]^\circ$) computed with and without extracted kernels. It demonstrates the importance of using the extracted kernel in the computation of the solenoidal part of the RHS ($\mathbf{P}^{\Lambda H}\mathbf{E}_{\text{ext}}$) and in the radiation of the solenoidal part of the current $\mathbf{J}_{\Lambda\mathbf{H}}$. It is clear that both are mandatory at very low frequency to match with the analytic solution (Mie series).

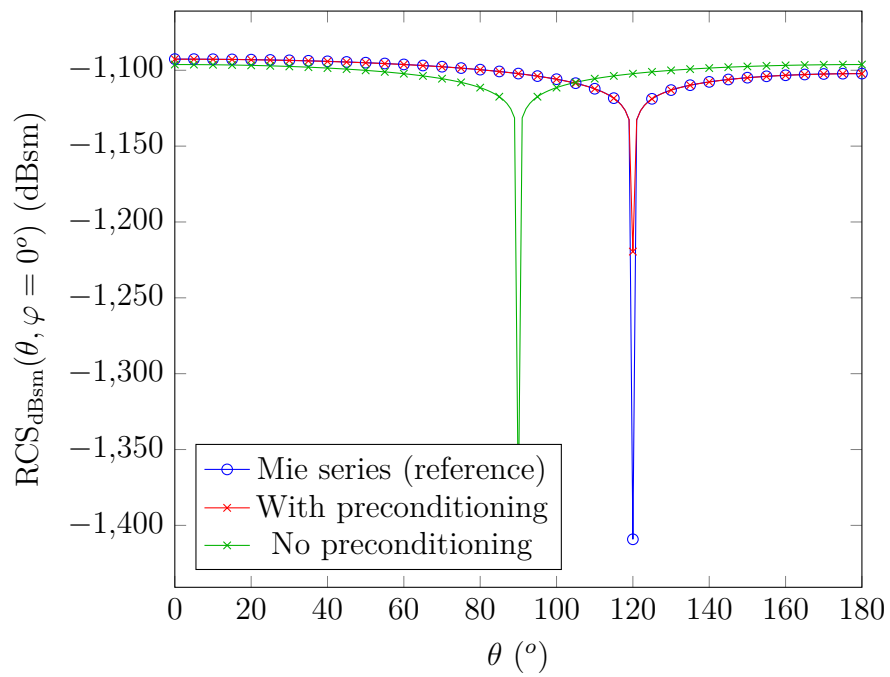


Figure 1.22: RCS of a PEC unit sphere with and without using the extracted kernel in the computation of the RHS and far field at $f = 10^{-20}$ Hz

Chapter 2

Low frequency and dense mesh stable IBC-EFIE

In this chapter, a formulation for the Impedance Boundary Condition - Electric Field Integral Equation (IBC-EFIE) is presented. It does not suffer from the dense mesh breakdown nor the low frequency breakdown. The stabilization of the formulation leverages the quasi-Helmholtz projectors.

2.1 Introduction

The Perfect Electric Conductor - Electric Field Integral Equation (PEC-EFIE) is often used to model accurately the scattering of time harmonic electromagnetic fields by a highly conducting scatterer. It is particularly efficient to simulate metals, but it exists a wide range of non-penetrable objects or materials that cannot fit in the PEC model. The Impedance Boundary Conditions (IBCs) can describe a wider range of non-penetrable objects. In fact, the PEC boundary condition is a particular case of IBC with an impedance that is equal to zero. The PEC boundary condition states that the only unknown is the tangential trace of the magnetic field or equivalently the surface electric current density $\mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H}$ (see (1.58)), and that the tangential trace of the electric field is zero on the boundary, or in terms of surface magnetic current density $\mathbf{M}_s = \mathbf{E} \times \hat{\mathbf{n}} = \mathbf{0}$. On the contrary, the IBC has both traces of the electric and magnetic fields as unknowns. At each point \mathbf{r} on the surface of the scatterer, they are coupled by a proportionality relation whose factor is the surface impedance $z^{\text{imp}}(\mathbf{r})$. The most simple IBC is the Leontovich IBC [13]

$$\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}(\mathbf{r}) = z^{\text{imp}}(\mathbf{r}) \hat{\mathbf{n}}(\mathbf{r}) \times (\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}(\mathbf{r})) \quad (2.1)$$

or equivalently, in terms of the surface currents

$$\mathbf{M}_s(\mathbf{r}) = -z^{\text{imp}}(\mathbf{r}) \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{J}_s(\mathbf{r}). \quad (2.2)$$

A common point between the PEC condition and the IBC is that they do not require the description of the domain in the interior of the scatterer (1.75) to model the

exterior domain. This is done by enforcing the condition between the surface currents instead. IBCs in electromagnetics have been the topic of several contributions [32], including generalized IBC [33], Combined Field Integral Equation (CFIE) [34], and self-dual formulations [35, 36].

The combination of the IBC together with the EFIE is the so-called IBC-EFIE [37]. Compared to the PEC-EFIE, the IBC-EFIE includes an additional term that is similar to the operator of the Magnetic Field Integral Equation (MFIE). And similarly to the PEC-EFIE, the IBC-EFIE is also plagued by several breakdowns including the low frequency breakdown and the dense mesh breakdown. So far, all the IBC formulations in electromagnetics suffer from at least one of these two ill-conditionings. The goal of this chapter is to present a formulation for the IBC-EFIE that is both stable at low frequency and for dense meshes.

The quasi-Helmholtz projectors have been recently introduced in the stabilization of the PEC-EFIE [38] to make it stable and accurate until arbitrary low frequency. Combined with a Calderón preconditioning, it results in a formulation whose resulting MoM system has a conditioning independent of the mesh size. However, the strategy used in the preconditioning of the PEC-EFIE cannot be straightforwardly applied to the IBC-EFIE due to the presence of the impedance and the MFIE operator.

The standard discretization of the IBC-EFIE is described in the first section. Then, an analysis of the IBC-EFIE operator is done using Vector Spherical Harmonics, on one hand at low frequency and on the other hand for dense discretizations. From the analysis, a preconditioner is built to stabilize the formulation. Numerical results are presented in the last section.

The content of this chapter has been published in a journal [39] and presented in a conference [40].

2.2 Impedance Boundary Condition - Electric Field Integral Equation

2.2.1 Integral equation of the IBC-EFIE

A scatterer with a boundary Γ is placed in the background medium whose characteristic impedance is η and the wave number is k . When the scatterer is illuminated by an incident electric field \mathbf{E}^{inc} , the surface currents \mathbf{J}_s and \mathbf{M}_s verify the EFIE (1.73a)

$$\eta \mathcal{T} \mathbf{J}_s + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{M}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \quad (2.3)$$

where the operators \mathcal{T} and \mathcal{K} are defined in (1.68a) and (1.69). The EFIE is completed with the Impedance Boundary Condition that is assumed to hold on the surface at every point $\mathbf{r} \in \Gamma$

$$\mathbf{M}_s(\mathbf{r}) = -z^{\text{imp}} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{J}_s(\mathbf{r}). \quad (2.4)$$

The impedance z^{imp} is a complex scalar that is assumed to be the uniformly the same on Γ . Depending on the model chosen for the impedance, z^{imp} can depend on

the frequency. The EFIE (2.3) together with the IBC (2.4) uniquely determine the solution currents \mathbf{J}_s and \mathbf{M}_s . Inserting (2.4) into (2.3) yields the IBC-EFIE

$$\eta\mathcal{T}\mathbf{J}_s - z^{\text{imp}}\left(\frac{\mathcal{I}}{2} - \mathcal{K}\right)(\hat{\mathbf{n}} \times \mathbf{J}_s) = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \quad (2.5)$$

that is to be solved in \mathbf{J}_s . In the following, the IBC-EFIE operator is defined as

$$\mathcal{S} = \eta\mathcal{T} - z^{\text{imp}}\left(\frac{\mathcal{I}}{2} - \mathcal{K}\right)(\hat{\mathbf{n}} \times \mathcal{I}) \quad (2.6)$$

such that the IBC-EFIE is conveniently rewritten as

$$\mathcal{S}\mathbf{J}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \quad (2.7a)$$

$$\mathbf{M}_s = -z^{\text{imp}}\hat{\mathbf{n}} \times \mathbf{J}_s. \quad (2.7b)$$

2.2.2 Discretization of the IBC-EFIE

The surface Γ is approximated by a triangular mesh on which N_s RWG basis functions $(\mathbf{f}_n)_{n=1}^{N_s}$ and BC basis functions $(\mathbf{g}_n)_{n=1}^{N_s}$ are constructed as defined in (1.90) and (1.92).

The RWG basis functions discretize the electric current density \mathbf{J}_s and the BC basis functions discretize the magnetic current density \mathbf{M}_s . These choices for the basis functions enable the construction of a conforming discretization as pointed out in [34]. Explicitly, the currents are expanded as

$$\mathbf{J}_s \approx \sum_{n=1}^{N_s} [\mathbf{J}]_n \mathbf{f}_n \quad (2.8a)$$

$$\mathbf{M}_s \approx \sum_{n=1}^{N_s} [\mathbf{M}]_n \mathbf{g}_n \quad (2.8b)$$

where \mathbf{J} and $\mathbf{M} \in \mathbb{C}^{N_s}$ are the vectors that contain the coefficients of the expansions. Both the EFIE (2.3) and the IBC (2.4) are tested with the rotated RWG functions $\hat{\mathbf{n}} \times \mathbf{f}_n$ ($n \in [1, N_s]$) which results in the following system of equations

$$\eta\mathbf{T}\mathbf{J} + \left(\frac{1}{2}\mathbf{G}_m - \mathbf{K}\right)\mathbf{M} = \mathbf{E} \quad (2.9a)$$

$$\mathbf{G}_m\mathbf{M} = -z^{\text{imp}}\mathbf{G}\mathbf{J} \quad (2.9b)$$

where

$$[\mathbf{T}]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}\mathbf{f}_n \rangle \quad (2.10a)$$

$$[\mathbf{K}]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{K}\mathbf{g}_n \rangle \quad (2.10b)$$

$$[\mathbf{G}_m]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathbf{g}_n \rangle \quad (2.10c)$$

$$[\mathbf{G}]_{mn} = \langle \mathbf{f}_m, \mathbf{f}_n \rangle \quad (2.10d)$$

$$[\mathbf{E}]_m = -\langle \mathbf{f}_m, \mathbf{E}^{\text{inc}} \rangle. \quad (2.10e)$$

In (2.9), the magnetic current \mathbf{M} can be substituted in the EFIE using the IBC to obtain the discrete counterpart of (2.5)

$$\eta \mathbf{T} \mathbf{J} - z^{\text{imp}} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K} \right) \mathbf{G}_m^{-1} \mathbf{G} \mathbf{J} = \mathbf{E} \quad (2.11a)$$

$$\mathbf{M} = -z^{\text{imp}} \mathbf{G}_m^{-1} \mathbf{G} \mathbf{J}. \quad (2.11b)$$

Regarding the previous equation, the discretization of the IBC-EFIE operator \mathcal{S} (2.6) is

$$\mathbf{S} = \eta \mathbf{T} - z^{\text{imp}} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K} \right) \mathbf{G}_m^{-1} \mathbf{G}. \quad (2.12)$$

The linear system corresponding to the MoM discretization of the IBC-EFIE is

$$\mathbf{S} \mathbf{J} = \mathbf{E}. \quad (2.13)$$

Unfortunately, this system is ill-conditioned. In particular, it suffers from the low-frequency breakdown and the dense mesh breakdown.

2.3 Analysis of the IBC-EFIE operator

2.3.1 Spectrum of the IBC-EFIE operator

In this section, the behaviour of the IBC-EFIE operator is studied using Vector Spherical Harmonics (VSH) on a sphere of radius a . The VSH \mathbf{X}_{lm} and \mathbf{U}_{lm} are defined in (1.130). It is reminded that \mathbf{X}_{lm} is solenoidal and that \mathbf{U}_{lm} is non-solenoidal. Using (1.131) and (1.132), the VSH expansion of the operator \mathcal{S} (2.6) is

$$\mathcal{S} \mathbf{X}_{lm} = \left(-\eta H_l^{(2)}(ka) - z^{\text{imp}} i H_l^{(2)'}(ka) \right) J_l(ka) \mathbf{U}_{lm} \quad (2.14a)$$

$$\mathcal{S} \mathbf{U}_{lm} = \left(\eta H_l^{(2)'}(ka) - z^{\text{imp}} i H_l^{(2)}(ka) \right) J_l'(ka) \mathbf{X}_{lm}. \quad (2.14b)$$

The singular values of the operator \mathcal{S} are the square roots of the eigenvalues of the operator $\mathcal{S}^* \mathcal{S}$. Using (1.133), the adjoint of \mathcal{S} is

$$\mathcal{S}^* = \eta \mathcal{T}^* + \overline{z^{\text{imp}}} \hat{\mathbf{n}} \times \left(\frac{\mathcal{I}}{2} - K^* \right). \quad (2.15)$$

Then, the eigenvalues of $\mathcal{S}^* \mathcal{S}$ are given by the following expressions

$$\mathcal{S}^* \mathcal{S} \mathbf{X}_{lm} = \left| -\eta H_l^{(2)}(ka) - z^{\text{imp}} i H_l^{(2)'}(ka) \right|^2 J_l(ka)^2 \mathbf{X}_{lm} \quad (2.16a)$$

$$\mathcal{S}^* \mathcal{S} \mathbf{U}_{lm} = \left| \eta H_l^{(2)'}(ka) - z^{\text{imp}} i H_l^{(2)}(ka) \right|^2 J_l'(ka)^2 \mathbf{U}_{lm} \quad (2.16b)$$

and consequently, the singular values of \mathcal{S} are

$$\sigma_{\mathbf{X},l} = \left| -\eta H_l^{(2)}(ka) - z^{\text{imp}} i H_l^{(2)'}(ka) \right| |J_l(ka)| \quad (2.17a)$$

$$\sigma_{\mathbf{U},l} = \left| \eta H_l^{(2)'}(ka) - z^{\text{imp}} i H_l^{(2)}(ka) \right| |J_l'(ka)| \quad (2.17b)$$

for all $l \geq 1$ and with a multiplicity $2l + 1$ because $m \in [-l, l]$. It must be noted that there are particular impedances for which the IBC-EFIE is resonant i.e. it exists impedances z^{imp} such that the singular values $\sigma_{\mathbf{X},l}$ and $\sigma_{\mathbf{U},l}$ are 0. As a result, the condition number of \mathbf{S} is unbounded for these impedances

$$\sigma_{\mathbf{X},l} = 0 \text{ if } z^{\text{imp}} = \eta i \frac{H_l^{(2)}(ka)}{H_l^{(2)'}(ka)} \quad (2.18a)$$

$$\sigma_{\mathbf{U},l} = 0 \text{ if } z^{\text{imp}} = -\eta i \frac{H_l^{(2)'}(ka)}{H_l^{(2)}(ka)}. \quad (2.18b)$$

Also, the IBC-EFIE has resonances for particular wave numbers k at high frequency due to the zeros of the Riccati-Bessel functions J_l and their derivatives J_l' . They are actually the same as the PEC-EFIE (see section 1.5.2.4).

2.3.2 Analytic solution of the IBC-EFIE on a sphere

In this section, the analytic solution of the IBC-EFIE is computed. The impedance z^{imp} and the frequency f are assumed to be different than their values that cancel $\langle \mathbf{X}_{lm}, \mathcal{S}\mathbf{U}_{lm} \rangle$ or $\langle \mathbf{U}_{lm}, \mathcal{S}\mathbf{X}_{lm} \rangle$. Using the orthonormality of the VSH, the currents can be expanded as

$$\mathbf{J}_s = \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle \mathbf{X}_{lm} + \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle \mathbf{U}_{lm} \quad (2.19a)$$

$$\mathbf{M}_s = \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{X}_{lm}, \mathbf{M}_s \rangle \mathbf{X}_{lm} + \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{U}_{lm}, \mathbf{M}_s \rangle \mathbf{U}_{lm}. \quad (2.19b)$$

The analytic solution is found by computing each inner product in the above equation. The solution for the electric current \mathbf{J}_s can be found by testing the IBC-EFIE (2.7a) with the VSH

$$\langle \mathbf{X}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle = \langle \mathbf{X}_{lm}, \mathcal{S}\mathbf{J}_s \rangle = \langle \mathbf{X}_{lm}, \mathcal{S}\mathbf{U}_{lm} \rangle \langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle \quad (2.20a)$$

$$\langle \mathbf{U}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle = \langle \mathbf{U}_{lm}, \mathcal{S}\mathbf{J}_s \rangle = \langle \mathbf{U}_{lm}, \mathcal{S}\mathbf{X}_{lm} \rangle \langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle. \quad (2.20b)$$

So,

$$\langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle = \frac{\langle \mathbf{U}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle}{\langle \mathbf{U}_{lm}, \mathcal{S}\mathbf{X}_{lm} \rangle} \quad (2.21a)$$

$$\langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle = \frac{\langle \mathbf{X}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle}{\langle \mathbf{X}_{lm}, \mathcal{S}\mathbf{U}_{lm} \rangle}. \quad (2.21b)$$

Then using the IBC (2.4), the inner products for the magnetic current are computed

$$\langle \mathbf{X}_{lm}, \mathbf{M}_s \rangle = z^{\text{imp}} \langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle \quad (2.22a)$$

$$\langle \mathbf{U}_{lm}, \mathbf{M}_s \rangle = -z^{\text{imp}} \langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle. \quad (2.22b)$$

In practice, \mathbf{E}^{inc} is chosen to be a plane wave (1.42) whose polarization $\hat{\mathbf{p}}$ and wave vector \mathbf{k} are aligned on coordinates axis so that the computation is more convenient. In the following, the analytic solutions are computed for the plane wave whose polarization is $\hat{\mathbf{p}} = \hat{\mathbf{x}}$ and whose wave vector is $\mathbf{k} = -k\hat{\mathbf{z}}$. It corresponds to an excitation in the form (see [41])

$$\mathbf{E}^{\text{inc}}(\mathbf{r}) = E_0 e^{ik\hat{\mathbf{z}} \cdot \mathbf{r}} \hat{\mathbf{x}} \quad (2.23a)$$

$$= 2\pi E_0 \sum_{l=1}^{\infty} i^l \sqrt{\frac{2l+1}{4\pi}} \left(j_l(ka) (\mathbf{X}_{l,-1} + \mathbf{X}_{l,1}) - \frac{1}{k} \nabla \times (j_l(ka) (\mathbf{X}_{l,-1} - \mathbf{X}_{l,1})) \right) \quad (2.23b)$$

$$= 2\pi E_0 \sum_{l=1}^{\infty} i^l \sqrt{\frac{2l+1}{4\pi}} \left(j_l(ka) (\mathbf{X}_{l,-1} + \mathbf{X}_{l,1}) - i\sqrt{l(l+1)} \frac{j_l(ka)}{ka} (\mathbf{Y}_{l,-1} - \mathbf{Y}_{l,1}) - \frac{J'_l(ka)}{ka} (\mathbf{U}_{l,-1} - \mathbf{U}_{l,1}) \right) \quad (2.23c)$$

where j_l are the spherical Bessel functions. The RHS is actually the tangential trace of the field. Taking the tangential trace of \mathbf{E}^{inc} eliminates the radial VSH components (\mathbf{Y}_{lm})

$$-\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}}(\mathbf{r}) = -2\pi E_0 \sum_{l=1}^{\infty} i^l \sqrt{\frac{2l+1}{4\pi}} \left(j_l(ka) (\mathbf{U}_{l,-1} + \mathbf{U}_{l,1}) + \frac{J'_l(ka)}{ka} (\mathbf{X}_{l,-1} - \mathbf{X}_{l,1}) \right) \quad (2.24a)$$

$$= \sum_{l=1}^{\infty} \sum_{m=\pm 1} \langle \mathbf{X}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle \mathbf{X}_{lm} + \sum_{l=1}^{\infty} \sum_{m=\pm 1} \langle \mathbf{U}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle \mathbf{U}_{lm} \quad (2.24b)$$

with

$$\langle \mathbf{X}_{l,\pm 1}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle = \pm 2\pi E_0 i^l \sqrt{\frac{2l+1}{4\pi}} \frac{J'_l(ka)}{ka} \quad (2.25a)$$

$$\langle \mathbf{U}_{l,\pm 1}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle = -2\pi E_0 i^l \sqrt{\frac{2l+1}{4\pi}} j_l(ka). \quad (2.25b)$$

The VSH expansion of the plane wave (2.25) together with the VSH expansion of the operator \mathcal{S} (2.14)

$$\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle = \left(-\eta H_l^{(2)}(ka) - z^{\text{imp}} i H_l^{(2)'}(ka) \right) J_l(ka) \quad (2.26a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{S} \mathbf{U}_{lm} \rangle = \left(\eta H_l^{(2)'}(ka) - z^{\text{imp}} i H_l^{(2)}(ka) \right) J'_l(ka) \quad (2.26b)$$

enable the computation of the VSH expansion of the current (2.21)

$$\langle \mathbf{X}_{l,\pm 1}, \mathbf{J}_s \rangle = \frac{E_0 i^l \sqrt{\pi(2l+1)}}{\left(\eta H_l^{(2)}(ka) + z^{\text{imp}} i H_l^{(2)'}(ka) \right) ka} \quad (2.27a)$$

$$\langle \mathbf{U}_{l,\pm 1}, \mathbf{J}_s \rangle = \frac{\pm E_0 i^l \sqrt{\pi(2l+1)}}{\left(\eta H_l^{(2)'}(ka) - z^{\text{imp}} i H_l^{(2)}(ka) \right) ka}. \quad (2.27b)$$

Finally, the VSH expansion for the magnetic current is computed using (2.22).

Usually, the quantity of interest are the scattered fields. The Mie theory describes a systematic method to obtain the series of coefficients (Mie series) of the VSH expansion of the fields from the boundary conditions. The Mie series for IBCs can be found in [19].

2.3.3 Low frequency behaviour

2.3.3.1 Low frequency breakdown

The asymptotic behaviour of the singular values of \mathcal{S} (2.17) is studied at low frequency by letting $k \rightarrow 0$. Also, it is assumed that z^{imp} is not near one of the resonant impedances (2.18), so that the equivalents below cannot be 0. Using the asymptotic behaviour of the Bessel family functions (1.136),

$$\sigma_{\mathbf{X},l} \underset{k \rightarrow 0}{\sim} \left| \frac{ika}{l} \eta + z^{\text{imp}} \right| \frac{l}{2l+1} \quad (2.28a)$$

$$\sigma_{\mathbf{U},l} \underset{k \rightarrow 0}{\sim} \left| \frac{l}{ika} \eta + z^{\text{imp}} \right| \frac{l+1}{2l+1}. \quad (2.28b)$$

It is assumed that $z^{\text{imp}}/\eta = o(1/(ka))$ when $k \rightarrow 0$ in the following which holds in all practical cases, so that z^{imp} can be neglected in $\sigma_{\mathbf{U},l}$ at low frequency. After the discretization, the order l of the harmonics is assumed to be truncated at l_{max} . So at low frequency the singular values of \mathbf{S} scale proportionality to

$$\sigma_{\mathbf{X},l_{\text{max}}} \underset{k \rightarrow 0}{\propto} \left| \frac{ika}{l_{\text{max}}} \eta + z^{\text{imp}} \right| \quad (2.29a)$$

$$\sigma_{\mathbf{U},l_{\text{max}}} \underset{k \rightarrow 0}{\propto} \frac{\eta l_{\text{max}}}{ka} \quad (2.29b)$$

and its condition number scales as

$$\text{cond}(\mathbf{S}) \underset{k \rightarrow 0}{\propto} \frac{\sigma_{\mathbf{U},l_{\text{max}}}}{\sigma_{\mathbf{X},l_{\text{max}}}} \underset{k \rightarrow 0}{\propto} \frac{\eta l_{\text{max}}}{ka \left| \frac{ika}{l_{\text{max}}} \eta + z^{\text{imp}} \right|}. \quad (2.30)$$

It must be noted that when $z^{\text{imp}} = 0$, the quadratic growing $1/k^2$ of the condition number of the PEC-EFIE (1.138) is retrieved. However if $|z^{\text{imp}}| \gg \eta ka$, the growing of the condition number is proportional to $1/(|z^{\text{imp}}| k)$.

2.3.3.2 Numerical cancellations

In addition to the ill-conditioning of the system at low frequency, there are also numerical cancellations on the RHS. For a plane wave \mathbf{E}^{inc} with a peak amplitude E_0 , the following scalings hold (2.25)

$$\langle \mathbf{X}_{lm}, \hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle \underset{k \rightarrow 0}{\propto} E_0 \quad (2.31a)$$

$$\langle \mathbf{U}_{lm}, \hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle \underset{k \rightarrow 0}{\propto} ikaE_0. \quad (2.31b)$$

In fact, the difference of scaling results in a loss of accuracy in the solenoidal part of the excitation field until the frequency is low enough to completely cancel it. These cancellations on the RHS are exactly the same as the PEC-EFIE (the RHS is independent of the boundary condition).

The scaling of the operator \mathcal{S} (2.26) at low frequency is

$$\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle \underset{k \rightarrow 0}{\propto} -\frac{ika}{l} \eta - z^{\text{imp}} \quad (2.32a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{S} \mathbf{U}_{lm} \rangle \underset{k \rightarrow 0}{\propto} \frac{l}{ika} \eta. \quad (2.32b)$$

Then, using the scaling of the RHS (2.31) and the LHS (2.32) in the expression for the currents (2.21) and (2.22) results in the following scalings

$$\langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle \underset{k \rightarrow 0}{\propto} \frac{ika}{\frac{ika}{l} \eta + z^{\text{imp}}} E_0 \quad (2.33a)$$

$$\langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle \underset{k \rightarrow 0}{\propto} \frac{ika}{\eta l} E_0 \quad (2.33b)$$

$$\langle \mathbf{X}_{lm}, \mathbf{M}_s \rangle \underset{k \rightarrow 0}{\propto} \frac{z^{\text{imp}}}{\eta} \frac{ika}{l} E_0 \quad (2.33c)$$

$$\langle \mathbf{U}_{lm}, \mathbf{M}_s \rangle \underset{k \rightarrow 0}{\propto} \frac{ika}{\frac{\eta}{z^{\text{imp}}} \frac{ika}{l} + 1} E_0. \quad (2.33d)$$

When $|z^{\text{imp}}| \ll \eta ka$ and in particular if $z^{\text{imp}} = 0$, there is a cancellation of the non-solenoidal part of the electric current and the solenoidal part of the magnetic current in the IBC-EFIE. There is a factor in the order of magnitude of k between the two parts of the currents. That is the same magnitude for the cancellation that occurs in the current of the PEC-EFIE. However, when $|z^{\text{imp}}| \gg \eta ka$, the factor between the two parts of the currents is in the order of z^{imp}/η . It can still lead to numerical cancellations depending on how the impedance z^{imp} varies with the frequency but the problem is less severe compared to the PEC-EFIE. Nevertheless it is still a potential problem that needs to be addressed.

2.3.4 Analysis of the dense mesh breakdown

The behaviour of the operator is again studied on a sphere of radius a with the VSH. As it is explained in the section 1.5.2.3, the dense mesh breakdown can be studied by

letting the order of the VSH l go to infinity. After the discretization, the average edge length of the discrete mesh Γ is h , and the spectrum of the operators are truncated to $l_{\max} \propto a/h$ (1.144). So the behaviour of the singular values of \mathcal{S} (2.17) when $l \rightarrow +\infty$ can be used to get an estimate of the condition number of \mathbf{S} when $h \rightarrow 0$. Using the asymptotic behaviour of the Bessel family functions (1.139) in (2.17) results in

$$\sigma_{\mathbf{X},l} \underset{l \rightarrow +\infty}{\sim} \begin{cases} \frac{|z^{\text{imp}}|}{2} & \text{if } z^{\text{imp}} \neq 0 \\ \frac{\eta ka}{2l} & \text{if } z^{\text{imp}} = 0 \end{cases} \quad (2.34a)$$

$$\sigma_{\mathbf{U},l} \underset{l \rightarrow +\infty}{\sim} \frac{\eta l}{2ka}. \quad (2.34b)$$

It can be read from (2.34a) that $\sigma_{\mathbf{X},l}$ behaves differently depending on whether $z^{\text{imp}} \neq 0$ or $z^{\text{imp}} = 0$. As the case $z^{\text{imp}} = 0$ corresponds to the PEC-EFIE, it won't be considered here because it is explained in the section 1.5.2.3 and the problem was solved in [38]. So in the following, $z^{\text{imp}} \neq 0$ is assumed and thus $\sigma_{\mathbf{X},l} \sim |z^{\text{imp}}|/2$ when $l \rightarrow +\infty$. Using the estimate $l_{\max} \propto a/h$, the condition number scales as

$$\text{cond}(\mathbf{S}) \underset{h \rightarrow 0}{\propto} \frac{\sigma_{\mathbf{U},l_{\max}}}{\sigma_{\mathbf{X},l_{\max}}} \underset{h \rightarrow 0}{\propto} \frac{\eta}{kh|z^{\text{imp}}|}. \quad (2.35)$$

Contrary to the PEC-EFIE for which the condition number grows quadratically with the inverse of the edge length h , in the IBC-EFIE it grows linearly. Indeed, the branch of the spectrum associated to the operator \mathcal{T}_s (whose singular values goes to 0 in the PEC-EFIE) is dominated by the identity (that is absent in the PEC-EFIE).

2.4 Solution of the low frequency problems

In this section, the low frequency breakdown and the numerical cancellations are addressed by leveraging the quasi-Helmholtz projectors (see section 1.5.3.3 for their introduction). In the RWG basis, \mathbf{P}^{AH} projects on the solenoidal subspace while \mathbf{P}^{Σ} projects on the non-solenoidal subspace. In the following, a left and a right preconditioners for the IBC-EFIE are built based on the analysis of the previous sections.

In the PEC-EFIE the left preconditioner (1.196) is designed to fix the numerical cancellation on the RHS by rescaling only the solenoidal part. The fact that this problem is exactly the same on the RHS of the IBC-EFIE (2.31) suggests the use of the same left preconditioner

$$\mathbf{L} = -\frac{1}{ika} \mathbf{P}^{\text{AH}} + \mathbf{P}^{\Sigma}. \quad (2.36)$$

Here the scalar a is no longer the radius of the sphere but a length parameter (ka has no unit) that should be in the order of magnitude of the diameter of the scatterer. By keeping a proportional to the diameter of the object, a global rescaling of the problem

(i.e. that rescales both of the geometry and the wavelength) lets ka invariant so it does not change the condition number of the preconditioned system.

The right preconditioner is designed to act on the electric current. The scaling of its non-solenoidal part (\mathbf{U}_{lm}) is roughly the same as in the PEC-EFIE, because the action of the impedance z^{imp} on this part was neglected by assuming $z^{\text{imp}}/\eta = o(1/(ka))$. Regarding the solenoidal part (\mathbf{X}_{lm}), it can be read from (2.33) that it scales like in the PEC-EFIE as E_0/η when $|z^{\text{imp}}| \ll \eta ka$, however it scales as kaE_0/z^{imp} when $|z^{\text{imp}}| \gg \eta ka$. Consequently, these two regimes are also present in the definition of the right preconditioner

$$\mathbf{R} = \frac{ika}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{P}^{\Lambda H} - ika \mathbf{P}^{\Sigma}. \quad (2.37)$$

The use of a right preconditioner introduces a new current-like unknown \mathbf{Y} that verifies

$$\mathbf{J} = \mathbf{R}\mathbf{Y}. \quad (2.38)$$

Using the above preconditioners in the IBC-EFIE (2.13) results in a low frequency stable IBC-EFIE

$$\mathbf{L}\mathbf{S}\mathbf{R}\mathbf{Y} = \mathbf{L}\mathbf{E}. \quad (2.39)$$

Once this auxiliary unknown \mathbf{Y} has been found by solving (2.39), the original current is retrieved using (2.38).

The preconditioned system can be analyzed using the quasi-Helmholtz decomposition of the system. The discretization of \mathcal{T} is split into its singular and hypersingular contributions as in (1.103)

$$\mathbf{T} = -ik\mathbf{T}_s + \frac{1}{ik}\mathbf{T}_h. \quad (2.40)$$

Also, the discretization of $(\frac{\mathcal{I}}{2} - \mathcal{K})(\hat{\mathbf{n}} \times \mathcal{I})$ is noted \mathbf{K}_+ such that

$$\mathbf{K}_+ = \left(\frac{1}{2}\mathbf{G}_m - \mathbf{K} \right) \mathbf{G}_m^{-1}\mathbf{G}. \quad (2.41)$$

Then, using the properties of \mathbf{T}_h (see (1.194)) to do some simplifications, the precon-

2. Low frequency and dense mesh stable IBC-EFIE

ditioned system is

$$\mathbf{LSR} = \left(-\frac{1}{ika} \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \right) \left(-ik\eta \mathbf{T}_s + \frac{\eta}{ik} \mathbf{T}_h - z^{\text{imp}} \mathbf{K}_+ \right) \left(\frac{ika}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{P}^{\Lambda H} - ika \mathbf{P}^{\Sigma} \right) \quad (2.42a)$$

$$= \mathbf{P}^{\Lambda H} \left(\frac{ik\eta}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{T}_s + \frac{z^{\text{imp}}}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{K}_+ \right) \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \left(\frac{k^2 a \eta}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{T}_s - \frac{ika z^{\text{imp}}}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{K}_+ \right) \mathbf{P}^{\Lambda H} \quad (2.42b)$$

$$+ \mathbf{P}^{\Lambda H} \left(-ik\eta \mathbf{T}_s - z^{\text{imp}} \mathbf{K}_+ \right) \mathbf{P}^{\Sigma} + \mathbf{P}^{\Sigma} \left(-k^2 a \eta \mathbf{T}_s - a \eta \mathbf{T}_h + ika z^{\text{imp}} \mathbf{K}_+ \right) \mathbf{P}^{\Sigma} = \mathbf{P}^{\Lambda H} \left(\frac{ik\eta}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{T}_s + \frac{z^{\text{imp}}}{ika + \frac{z^{\text{imp}}}{\eta}} \mathbf{K}_+ \right) \mathbf{P}^{\Lambda H} - a \eta \mathbf{P}^{\Sigma} \mathbf{T}_h \mathbf{P}^{\Sigma} - z^{\text{imp}} \mathbf{P}^{\Lambda H} \mathbf{K}_+ \mathbf{P}^{\Sigma} + \mathbf{O}_{k \rightarrow 0}(k). \quad (2.42c)$$

If $\lim_{k \rightarrow 0} z^{\text{imp}}$ is finite, which is the case in all relevant models for the impedance, then \mathbf{LSR} has a static limit (noted $\lim_{k \rightarrow 0} \mathbf{LSR}$). Thus the condition number of the preconditioned system is bounded and tends to $\text{cond}(\lim_{k \rightarrow 0} \mathbf{LSR})$. In other words, the preconditioned system is immune to the low frequency breakdown. In the case $|z^{\text{imp}}| \ll \eta ka$, the preconditioned IBC-EFIE (2.42) reduces to the preconditioned PEC-EFIE (1.198) with the following asymptotic behaviour

$$\mathbf{LSR} = \frac{\eta}{a} \mathbf{P}^{\Lambda H} \mathbf{T}_s \mathbf{P}^{\Lambda H} - a \eta \mathbf{P}^{\Sigma} \mathbf{T}_h \mathbf{P}^{\Sigma} + \mathbf{O}_{k \rightarrow 0}(k). \quad (2.43)$$

On the contrary, if $|z^{\text{imp}}| \gg \eta ka$, then (2.42) behaves as

$$\mathbf{LSR} = \eta \mathbf{P}^{\Lambda H} \mathbf{K}_+ \mathbf{P}^{\Lambda H} - a \eta \mathbf{P}^{\Sigma} \mathbf{T}_h \mathbf{P}^{\Sigma} - z^{\text{imp}} \mathbf{P}^{\Lambda H} \mathbf{K}_+ \mathbf{P}^{\Sigma} + \mathbf{O}_{k \rightarrow 0}(k). \quad (2.44)$$

Note that a part of the spectrum ($z^{\text{imp}} \mathbf{P}^{\Lambda H} \mathbf{K}_+ \mathbf{P}^{\Sigma}$) still depends on z^{imp} at low frequency. It is a problem if $|z^{\text{imp}}| \gg \eta$ because the condition number is proportional to $|z^{\text{imp}}|$ in this case. However, the case $|z^{\text{imp}}| \gg \eta$ can still be managed by considering the IBC-MFIE (1.73b)

$$-\frac{1}{\eta} \mathcal{T} \mathbf{M}_s + \frac{1}{z^{\text{imp}}} \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \hat{\mathbf{n}} \times \mathbf{M}_s = \hat{\mathbf{n}} \times \mathbf{H}^{\text{inc}} \quad (2.45a)$$

$$\mathbf{J}_s = \frac{1}{z^{\text{imp}}} \hat{\mathbf{n}} \times \mathbf{M}_s \quad (2.45b)$$

that has been rewritten such that the roles of \mathbf{M}_s and \mathbf{J}_s are exchanged compared to the IBC-EFIE. Also, the impedances are replaced by their respective admittances. By preconditioning the IBC-MFIE in an analogous fashion as the IBC-EFIE, the case $|z^{\text{imp}}| \gg \eta$ or equivalently $1/|z^{\text{imp}}| \ll 1/\eta$ is in fact not problematic because the part of the spectrum $(1/z^{\text{imp}}) \mathbf{P}^{\Lambda H} \mathbf{K}_+ \mathbf{P}^{\Sigma}$ is dominated by the rest of the spectrum that scales as $1/\eta$.

2.5 Solution of the dense mesh breakdown

The dense mesh breakdown is addressed with a Calderón-like preconditioning (i.e. preconditioning by a composition of operators). As it is shown in the previous section, in the IBC-EFIE only one part of the spectrum needs a regularization. The idea is to apply the operator \mathcal{T}_s only on the part of the spectrum (2.34b) that is unbounded. Again, the selection of this part of the spectrum is achieved by leveraging the quasi-Helmholtz projectors (see section 1.5.3.3). To combine the operators conformingly, the dual discretization of \mathcal{T}_s with BC basis functions \mathbf{g}_n is used. It corresponds to the matrix \mathbb{T}_s (1.109) whose entries are $[\mathbb{T}_s]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{g}_m, \mathcal{T}_s \mathbf{g}_n \rangle$. Similarly, the dual projectors \mathbf{P}^Λ and $\mathbf{P}^{\Sigma H}$ (1.199) are used. In the BC basis, \mathbf{P}^Λ projects on the non-solenoidal subspace while $\mathbf{P}^{\Sigma H}$ projects on the solenoidal subspace. Based on the previous considerations, the dense mesh preconditioner is defined as

$$\tilde{\mathbb{T}} = \frac{1}{a} \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{P}^{\Sigma H} + \mathbf{P}^\Lambda. \quad (2.46)$$

The link between the preconditioner that is discretized with BC functions and the IBC-EFIE that is discretized with RWG functions is ensured by the inverse of the mix-Gram matrix \mathbf{G}_m^{-1} . The IBC-EFIE that is low frequency stable and dense mesh stable is

$$\tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{S} \mathbf{R} \mathbf{Y} = \tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{E}. \quad (2.47)$$

2.6 Implementation details

To avoid the numerical cancellations in the computation of the RHS, the zeros in the computations must be enforced manually. In particular, for an incident plane wave $\mathbf{E}^{\text{inc}}(\mathbf{r}) = \mathbf{E}_0 e^{-i\mathbf{k} \cdot \mathbf{r}}$, the extracted kernel $e^{-i\mathbf{k} \cdot \mathbf{r}} - 1$ must be used instead of $e^{-i\mathbf{k} \cdot \mathbf{r}}$ in the computation of the solenoidal part of the RHS. The vectors \mathbf{E} and \mathbf{E}_{ext} are defined as

$$[\mathbf{E}]_m = - \langle \mathbf{f}_m, \mathbf{E}_0 e^{-i\mathbf{k} \cdot \mathbf{r}} \rangle \quad (2.48a)$$

$$[\mathbf{E}_{\text{ext}}]_m = - \langle \mathbf{f}_m, \mathbf{E}_0 (e^{-i\mathbf{k} \cdot \mathbf{r}} - 1) \rangle. \quad (2.48b)$$

Mathematically, $\mathbf{P}^{\Lambda H} \mathbf{E} = \mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}}$ holds, but numerically, only $\mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}}$ provides the correct result at low frequency because $\mathbf{P}^{\Lambda H} \mathbf{E}$ has numerical cancellations. The RHS is computed as

$$\tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{L} \mathbf{E} = \frac{1}{a} \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{G}_m^{-1} \mathbf{P}^{\Sigma} \mathbf{E} \quad (2.49a)$$

$$- \frac{1}{ika^2} \mathbf{P}^{\Sigma H} \mathbb{T}_s \mathbf{P}^{\Sigma H} \mathbf{G}_m^{-1} \mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}} \quad (2.49b)$$

$$- \frac{1}{ika} \mathbf{P}^\Lambda \mathbf{G}_m^{-1} \mathbf{P}^{\Lambda H} \mathbf{E}_{\text{ext}}. \quad (2.49c)$$

The LHS is also rewritten to remove explicitly the products that are $\mathbf{0}$

$$\tilde{\mathbb{T}}\mathbf{G}_m^{-1}\mathbf{L}\mathbf{S}\mathbf{R} = \frac{\eta}{a}\mathbf{P}^\Lambda\mathbf{G}_m^{-1}\mathbf{T}_s\mathbf{R} \quad (2.50a)$$

$$- \frac{ik\eta}{a}\mathbf{P}^{\Sigma H}\mathbb{T}_s\mathbf{P}^{\Sigma H}\mathbf{G}_m^{-1}\mathbf{L}\mathbf{T}_s\mathbf{R} \quad (2.50b)$$

$$- \eta\mathbf{P}^{\Sigma H}\mathbb{T}_s\mathbf{G}_m^{-1}\mathbf{T}_h \quad (2.50c)$$

$$- z^{\text{imp}}\tilde{\mathbb{T}}\mathbf{G}_m^{-1}\mathbf{L}\left(\frac{1}{2}\mathbf{G}_m - \mathbf{K}\right)\mathbf{G}_m^{-1}\mathbf{G}\mathbf{R}. \quad (2.50d)$$

Finally, when the unknown \mathbf{Y} is found, the currents \mathbf{J} and \mathbf{M} are retrieved using (2.38) and (2.11b). The solenoidal parts ($\mathbf{J}_{\Lambda H}$ and $\mathbf{M}_{\Sigma H}$) and non-solenoidal parts (\mathbf{J}_Σ and \mathbf{M}_Λ) of the current are

$$\mathbf{J}_{\Lambda H} = \frac{ika}{ika + \frac{z^{\text{imp}}}{\eta}}\mathbf{P}^{\Lambda H}\mathbf{Y} \quad (2.51a)$$

$$\mathbf{J}_\Sigma = -ika\mathbf{P}^{\Sigma}\mathbf{Y} \quad (2.51b)$$

$$\mathbf{M}_{J,\Lambda H} = -z^{\text{imp}}\mathbf{G}_m^{-1}\mathbf{G}\mathbf{J}_{\Lambda H} \quad (2.51c)$$

$$\mathbf{M}_{J,\Sigma} = -z^{\text{imp}}\mathbf{G}_m^{-1}\mathbf{G}\mathbf{J}_\Sigma \quad (2.51d)$$

$$\mathbf{M}_{\Sigma H} = \mathbf{P}^{\Sigma H}\mathbf{M}_{J,\Lambda H} + \mathbf{P}^{\Sigma H}\mathbf{M}_{J,\Sigma} \quad (2.51e)$$

$$\mathbf{M}_\Lambda = \mathbf{P}^\Lambda\mathbf{M}_{J,\Lambda H} + \mathbf{P}^\Lambda\mathbf{M}_{J,\Sigma}. \quad (2.51f)$$

Note the use of the intermediate quantities noted $\mathbf{M}_{J,\Lambda H}$ and $\mathbf{M}_{J,\Sigma}$ to get the two parts of the magnetic current. They verify (mathematically but not numerically) $\mathbf{J} = \mathbf{J}_{\Lambda H} + \mathbf{J}_\Sigma$ and $\mathbf{M} = \mathbf{M}_{\Sigma H} + \mathbf{M}_\Lambda$, but in order to avoid reintroducing numerical cancellations in the current, these components of the currents should be kept separately. The extracted kernel should be used instead of the usual kernel for the fields scattered by the solenoidal parts of the currents. In particular, the scattering of $\mathbf{J}_{\Lambda H}$ and $\mathbf{M}_{\Sigma H}$ in the far field is computed using the extracted kernel $e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}'} - 1$ instead of the regular kernel $e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}'}$ as in (1.188).

2.7 Numerical results

The formulation has been compared against several other formulations:

- "This work" is the low frequency and dense mesh stable IBC-EFIE that corresponds to (2.47),
- "IBC-EFIE" is the non-stabilized system in (2.13),
- "IBC-CFIE" corresponds to [34],
- "IBC self-dual" corresponds to [35],

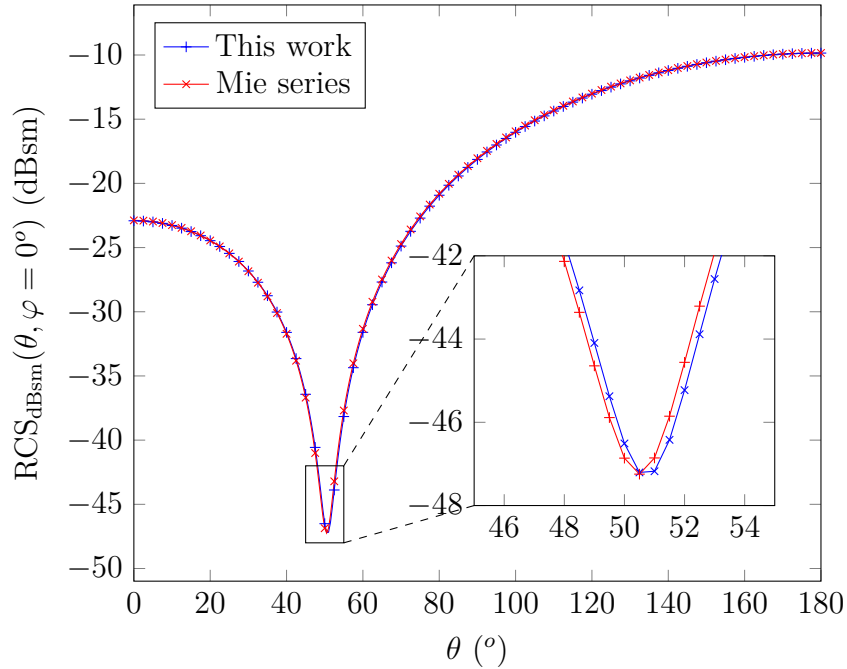


Figure 2.1: RCS of an IBC unit sphere ($f = 10^7$ Hz, $z^{\text{imp}} = (0.7 + 0.7i)\eta$, $h \approx 0.10$ m)

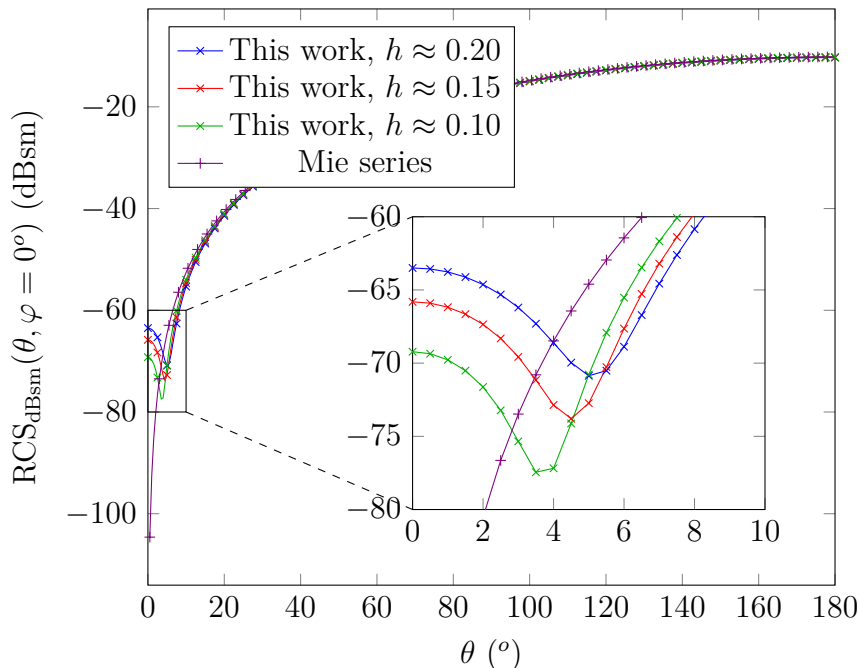
- "Loop-Star" is the low frequency stable system (2.39) that uses a Loop-Star decomposition [5] (see section 1.5.3.2) instead of a quasi-Helmholtz decomposition in \mathbf{L} and \mathbf{R} ,
- "Mie series" is the analytic solution on a sphere [19].

The length parameter a that is present in the preconditioners is set to 1 meter. Note that additional numerical results related to this work are also presented in the next chapter and in [39, 40].

2.7.1 Correctness of the formulation

The correctness of the formulation is first tested on a sphere that has an analytic solution available. The sphere radius is 1 meter. The excitation is a plane wave (1.42) with a direction of propagation $\hat{\mathbf{k}} = -\hat{\mathbf{z}}$, a polarization $\hat{\mathbf{p}} = \hat{\mathbf{x}}$ and a peak amplitude $E_0 = 1$ V/m. The RCS (1.47) is computed in the xz plane ($\varphi = 0^\circ$ and $\theta \in [0, 180]^\circ$) which corresponds to a so-called VV-polarization. The figure 2.1 ($f = 10^7$ Hz, $h \approx 0.1$ m) shows the solution for the impedance $z^{\text{imp}} = (0.7 + 0.7i)\eta$. It is clear that the solution that is obtained with this formulation converges to the analytic solution.

An issue where there is major discrepancy is when the back-scattered field vanishes. The figure 2.2 ($f = 10^7$ Hz) shows the solution for the particular impedance $z^{\text{imp}} = \eta$. In this case, the back-scattered field (at $\theta = 0^\circ$) is exactly $\mathbf{0}$ since the field scattered by \mathbf{J}_s is the opposite of the field scattered by \mathbf{M}_s . Unfortunately in


 Figure 2.2: RCS of an IBC unit sphere ($f = 10^7$ Hz, $z^{\text{imp}} = \eta$)

this formulation, the exact zero cannot be achieved numerically. Though, further increasing the discretization density of the mesh (as h decreases) shows that there is a convergence. This is likely to be due to the presence of numerical errors in the computation of the far field with the BC basis functions whose representation of the current is based on the barycentric refinement of the mesh, whereas the RWG functions are based on the primal mesh. Since the domains of numerical integration (and therefore the integration rules) of the currents are different, some discrepancies are expected. In fact, the IBC-CFIE in [34] that has been implemented as a reference also suffers from this problem since it has the same discretization for the currents (RWG basis functions for \mathbf{J}_s and BC basis functions for \mathbf{M}_s), however, the self-dual formulation [35] uses RWG functions for both \mathbf{J}_s and \mathbf{M}_s and has not this issue.

The solution found with this formulation has been compared to the solution obtained with the IBC-CFIE in [34] and the self-dual formulation in [35] on a unit sphere with $f = 10^7$ Hz, $z^{\text{imp}} = (0.7 + 0.7i)\eta$ and $h \approx 0.15$ m. The same integration rules have been used. The relative error on the solution \mathbf{J} of a formulation relative to the solution \mathbf{J}' of another formulation is computed as

$$\frac{\|\mathbf{J} - \mathbf{J}'\|}{\|\mathbf{J}\|}. \quad (2.52)$$

Compared to the IBC-CFIE in [34], the relative error is 17×10^{-4} on both \mathbf{J} and \mathbf{M} . Compared to the self-dual IBC formulation in [35], the relative error is 11×10^{-4} on \mathbf{J} (the error on \mathbf{M} is meaningless since it is a RWG expansion vector in [35] and a BC expansion vector in this work). As a comparison, the error on \mathbf{J} between the IBC-CFIE and the self-dual IBC is 8×10^{-4} . So it can be concluded that the 3

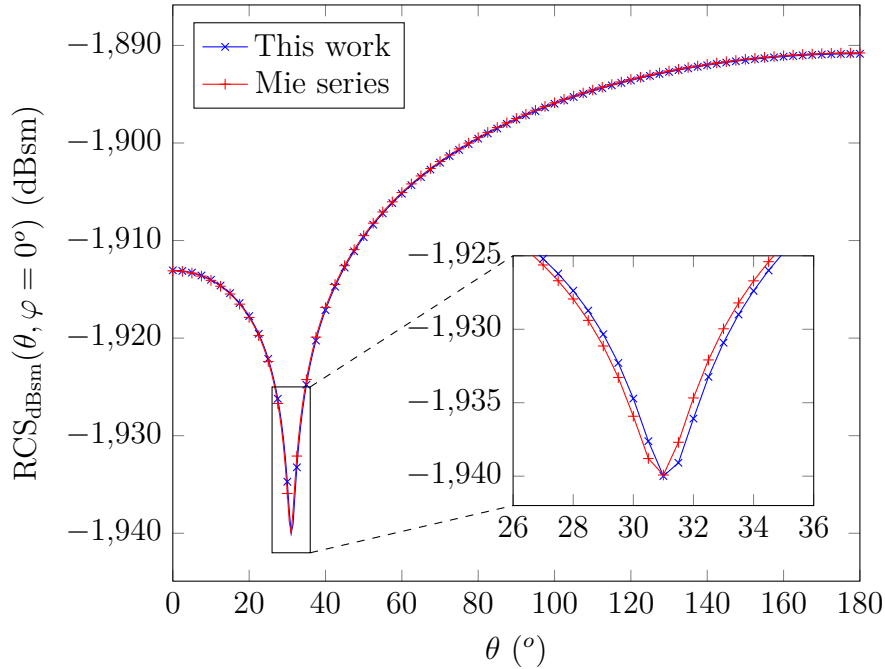


Figure 2.3: RCS of an IBC unit sphere ($f = 10^{-40}$ Hz, $z^{\text{imp}} = (0.1 + 2i) \times 10^{-47}\eta$, $h \approx 0.15$ m)

formulations converge to the same solution.

To ensure that the formulation is low frequency stable and that all the possible numerical cancellations have been properly addressed, the RCS is computed at a quasi static frequency ($f = 10^{-40}$ Hz, $z^{\text{imp}} = (0.1 + 2i) \times 10^{-47}\eta$, $h \approx 0.15$ m) in the figure 2.3. Again, even though the IBC is not valid at this frequency it pushes all the possible numerical problems to their limits to make sure they are all correctly addressed. The solution is recovered and the condition number of the preconditioned system is 6.61 which shows that the formulation does not suffers from the low frequency breakdown nor numerical cancellations.

2.7.2 Stability of the formulation

The figure 2.4 shows the condition numbers of the linear systems as functions of the frequency for several formulations. The impedance ($z^{\text{imp}} = (0.8 + 0.8i)\eta$) and the mesh parameter ($h \approx 0.15$ m) are kept constant. The growing of the condition number of the non-preconditioned IBC-EFIE is proportional to the inverse of the frequency. This matches with the predicted asymptotic behaviour (2.30). Also, it is clear that the formulation in this work is immune to the low frequency breakdown.

The formulation is also stable when the impedance is a function of the frequency. For conductors whose conductivity is σ' and whose permeability is μ' , most of the current flows in a thin layer on the boundary (the so-called skin, see section 1.4). In such a medium, the permittivity is approximately $\epsilon' \approx \sigma'/(i\omega)$. The associated impedance is $z^{\text{imp}} = \sqrt{\mu'/\epsilon'} \approx \sqrt{\frac{\mu'\omega}{2\sigma'}}(1 + i)$ i.e. the impedance varies proportionally

2. Low frequency and dense mesh stable IBC-EFIE

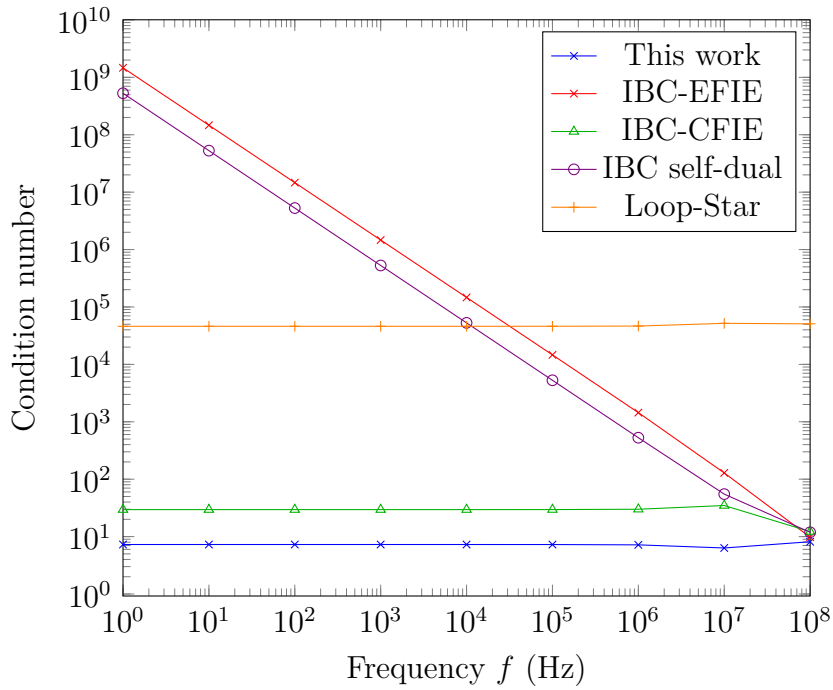


Figure 2.4: Condition number as a function of the frequency f

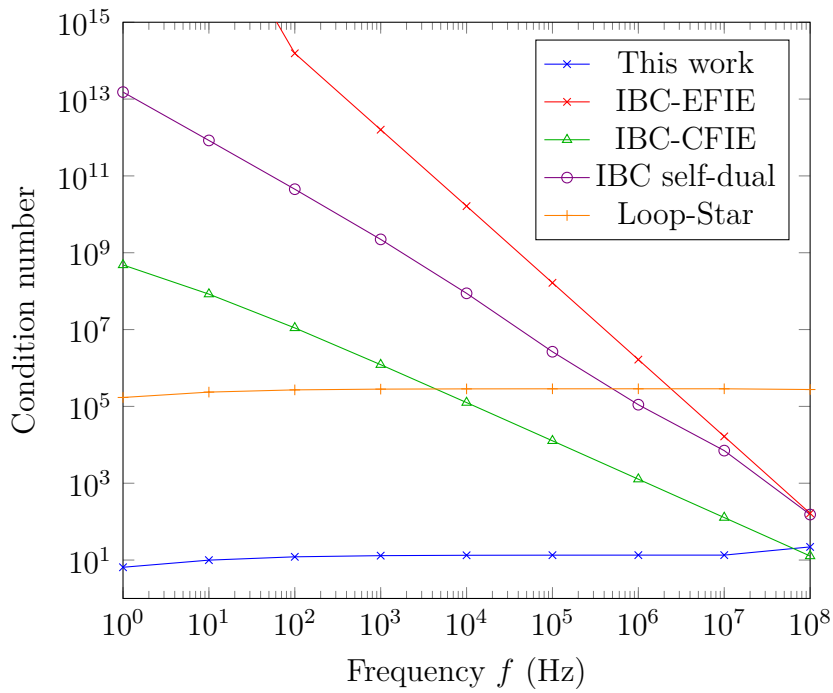


Figure 2.5: Condition number as a function of the frequency f when the impedance z^{imp} also varies with the frequency

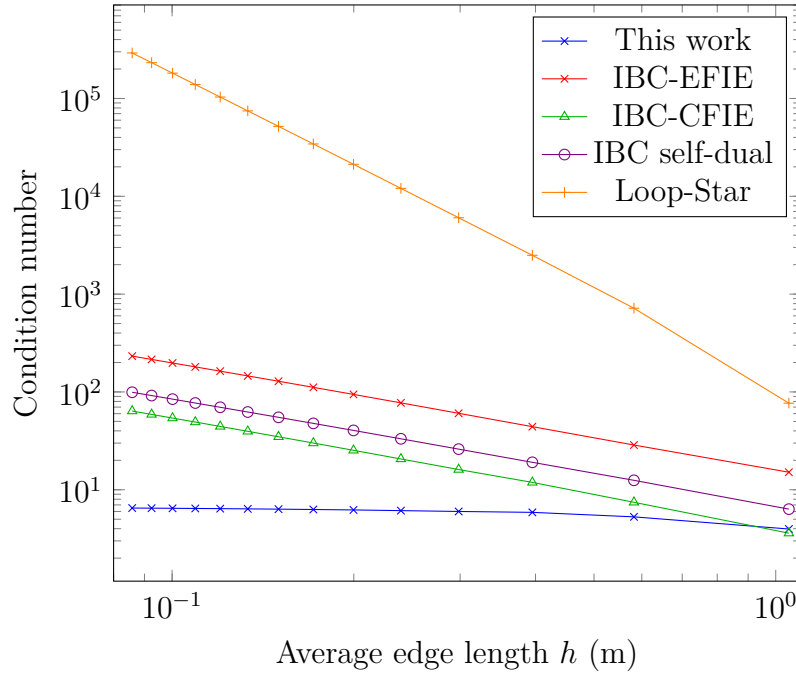


Figure 2.6: Condition number as a function of the average edge length h

to the square root of the frequency. In the figure 2.5, such a model for the impedance is used with the copper parameters ($\sigma' = 5.69 \times 10^7$ S/m, $\mu' = \mu_0$). The frequency is pushed beyond the validity of the IBC to study the behaviour of the formulations. As it can be observed, this work is immune to the low frequency breakdown even when z^{imp} is a function of the frequency f .

The behaviour of the formulation in the dense mesh regime is illustrated in the figure 2.6 where the condition number is a function of the average edge length h ($z^{\text{imp}} = (0.8 + 0.8i)\eta$ and $f = 10$ MHz are fixed). The condition number is stable for this work which demonstrates that it is immune from the dense mesh breakdown, contrary to the non-preconditioned formulations that have a growing of the condition number proportional to $1/h$.

2.7.3 Realistic case scenario

The formulation is used on a realistic example where the scatterer is a F-117 Nighthawk aircraft. The mesh is pictured in the figure 2.7. The surface electric current density is represented in the figure 2.8a and the magnetic current density is in the figure 2.8b when $z^{\text{imp}} = \eta$. The incident plane wave is vertically polarized and its angle of incidence is $\theta = 90^\circ$ (horizon plane) and $\varphi = 20^\circ$ at $f = 10$ MHz. The norms of the real parts of the currents have been evaluated in the center of each faces and then interpolated to obtain a smooth visualization. Regarding the effect of the preconditioners

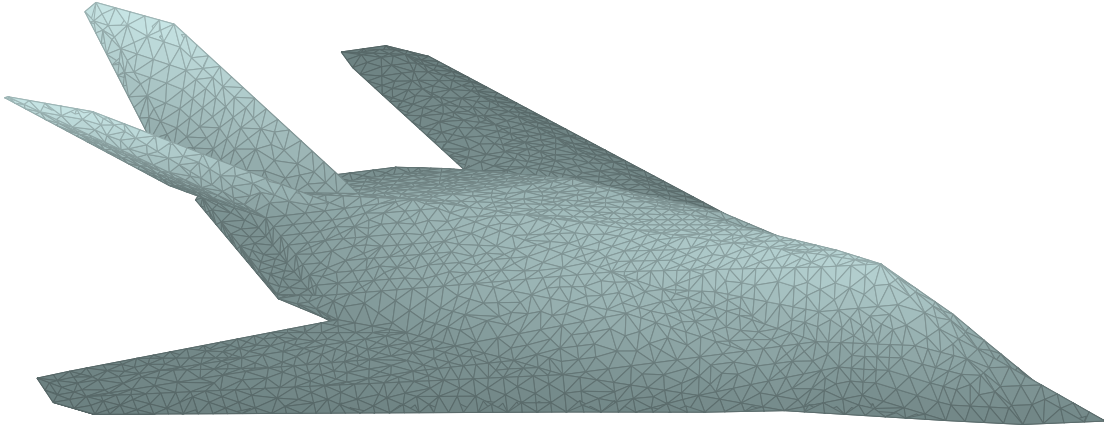


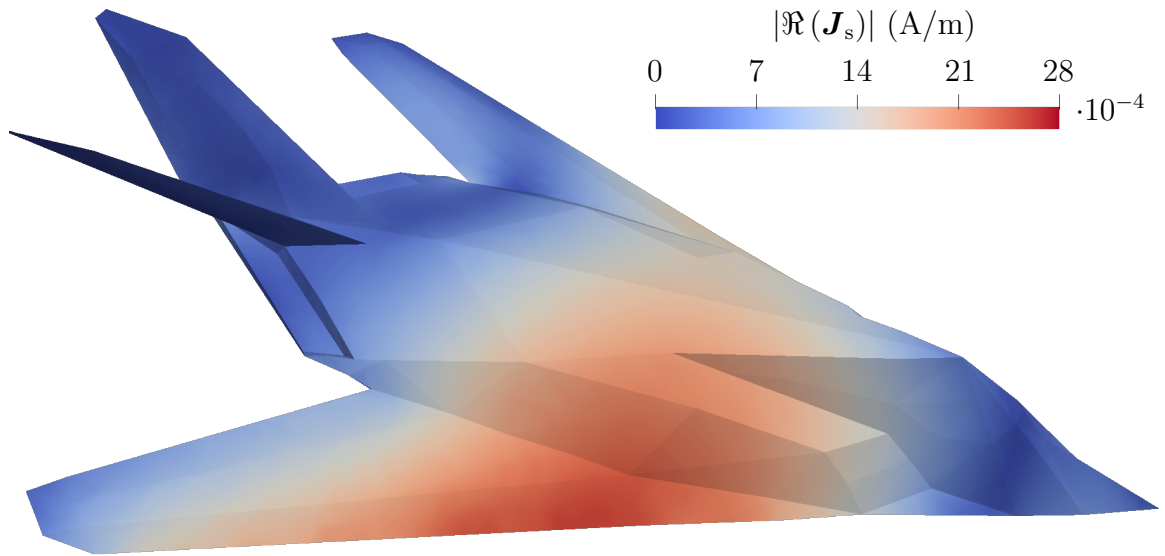
Figure 2.7: Aircraft mesh ($N_s = 11199$, $h \approx 0.29$ m)

2.8 Conclusion and future work

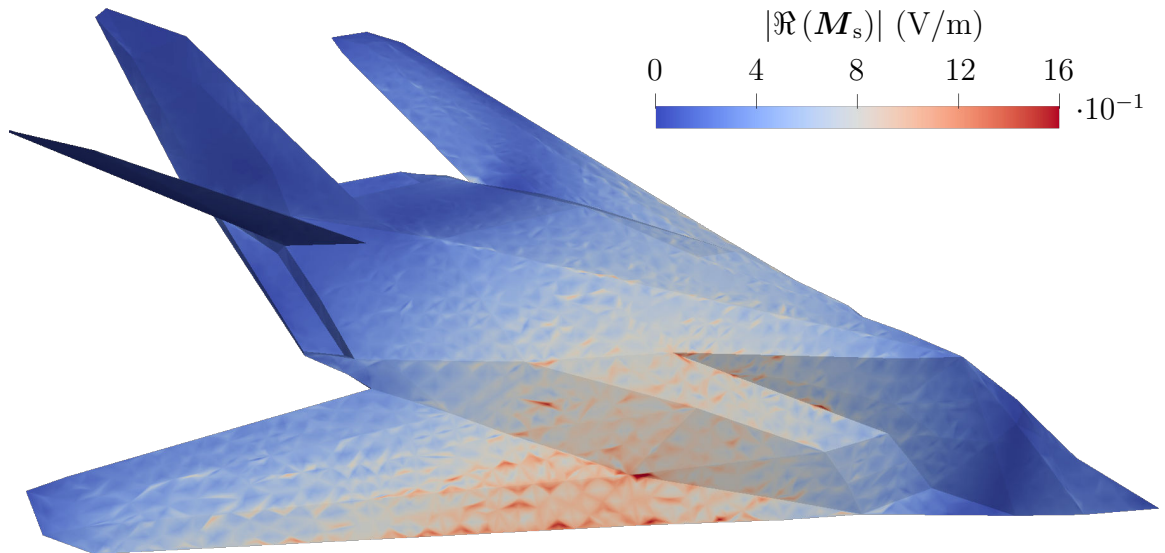
In this work, an IBC-EFIE that is stable for arbitrarily low frequency and dense meshes has been presented. On one hand, the low frequency stabilization involves quasi-Helmholtz projectors to independently rescale in frequency the Helmholtz components of the incident field and the current. On the other hand, the dense mesh stabilization involves a kind of Calderón preconditioning that regularizes only the unbounded branch of the spectrum, while letting unchanged the other part that is already dominated by the identity present in the IBC-EFIE.

There is a lot of room for future investigations. Here is a non-exhaustive list of known issues or possible axes of research:

- Non-uniform impedance z^{imp} . In this work z^{imp} is the same everywhere on the mesh, and an obvious generalization is to have an impedance that is not constant in space.
- Anisotropic impedance. Investigations can be done where the impedance is a tensor. In this case the IBC is (1.116a).
- Spurious resonance in frequency. In this work, the formulation is not free from high frequency resonances. This should be solved with an IBC-CFIE.
- Zero impedance. In the limit $z^{\text{imp}} \rightarrow 0$, this formulation differs from the classical PEC-EFIE for the dense mesh preconditioning. As a result, a dense mesh breakdown is effectively present. A simple but inelegant solution is to use the preconditioner of the PEC-EFIE (1.202) instead of (2.46) in these cases.
- Back scattering at $z^{\text{imp}} = \eta$. The figure 2.2 shows that on a sphere at $z^{\text{imp}} = \eta$ there is no back scattering, but the numerical results show a discrepancy with the analytic solution.
- High impedance. It can be read from (2.44) that a part of the spectrum of the preconditioned system depends on z^{imp} and effectively results in an ill-



(a) Norm of the real part of the surface electric current density $|\Re(\mathbf{J}_s)|$



(b) Norm of the real part of the surface magnetic current density $|\Re(\mathbf{M}_s)|$

Figure 2.8: Surface current densities on the aircraft at $f = 10$ MHz with $z^{\text{imp}} = \eta$

2. *Low frequency and dense mesh stable IBC-EFIE*

conditioned system if $|z^{\text{imp}}| \gg \eta$. As mentioned in (2.45), this issue can be alleviated using the IBC-MFIE.

Chapter 3

Multiplicative preconditioner for a new IBC-EFIE

In this chapter, a new Impedance Boundary Condition is introduced. It enables the construction of a multiplicative preconditioner for the IBC-EFIE that reduces to the classical multiplicative Calderón preconditioner of the PEC-EFIE [7] in the limit $z^{\text{imp}} \rightarrow 0$. This results in an IBC-EFIE formulation that is low frequency and dense mesh stable.

3.1 Introduction

In the previous chapter, a low frequency and dense mesh stable IBC-EFIE has been developed by leveraging a quasi-Helmholtz decomposition with a rescaling in frequency and a Calderón-like preconditioning. There are several considerations on the formulation of the previous chapter that have motivated the work presented in this chapter.

- When the impedance z^{imp} is equal to 0, the previous formulation does not reduce to the classical well-conditioned PEC-EFIE [7]. In particular, it still suffers from the dense mesh breakdown.
- The quasi-Helmholtz projectors are useful to obtain the low frequency stability but they make the formulation more complicated. More practically, they are not trivial to implement.
- The previous formulation enables an accurate computation of the currents down to arbitrary low frequencies that are far beyond the range of validity of IBC models.

These considerations are not really problematic, but they are enough to think about an alternative formulation that could take advantage of these. In particular, this chapter is dedicated to the development of an IBC-EFIE that is dense mesh stable even when $z^{\text{imp}} = 0$, that does not use quasi-Helmholtz projectors, and that is stable in the range of frequency where the IBC are relevant, but not arbitrary low frequency

3. Multiplicative preconditioner for a new IBC-EFIE

as numerical cancellations would still occur. To obtain a complete low frequency stability (i.e. free of numerical cancellations and with a stable condition number till arbitrary low frequency), the quasi-Helmholtz projectors would still be required.

The most widespread IBC is the Leontovich IBC that reads

$$\mathbf{M}_s = -z^{\text{imp}} \hat{\mathbf{n}} \times \mathbf{J}_s. \quad (3.1)$$

This relation between the tangential traces of the electric and magnetic field actually comes from the study of the transmission of the fields at an infinite planar interface. This IBC is a valid model for sufficiently smooth and conducting scatterers for which the fields interactions are concentrated on the boundary. In this work, a new IBC is introduced

$$\mathbf{M}_s = -z^{\text{imp}} \mathcal{N}_\delta \mathbf{J}_s. \quad (3.2)$$

where \mathcal{N}_δ is an operator that depends on a length δ . δ doesn't necessarily correspond to the skin depth. Instead, it should be chosen to leave the solution of the boundary problem unchanged, and at the same time to improve on the computational properties of the IBC-EFIE which enables the preconditioning. In particular, it enables the construction of a multiplicative preconditioner that reduces to the classical Calderón preconditioner for the PEC-EFIE [7] when $z^{\text{imp}} = 0$.

The new IBC is first introduced. It is used to build an IBC-EFIE. Then, the properties of the new IBC are studied in a Vector Spherical Harmonics (VSH) basis. This enables building an analytic solution for this new IBC-EFIE. Then, the asymptotic behaviour of the formulation is analyzed at low frequency and for dense meshes. Based on this analysis, a Calderón multiplicative preconditioner can be constructed to stabilize the formulation. Finally, numerical results are presented.

Preliminary work on this new IBC has been presented at a conference [42].

3.2 Background and notations

3.2.1 Construction of the IBC-EFIE

The background medium has a permittivity ε , a permeability μ and a characteristic impedance $\eta = \sqrt{\mu/\varepsilon}$. The angular frequency is ω and the wave number is $k = \omega\sqrt{\mu\varepsilon}$. The scatterer with a boundary noted Γ is placed in the background medium and is illuminated by an incident electric field \mathbf{E}^{inc} . This excitation induces the surface currents \mathbf{J}_s and \mathbf{M}_s that verify the EFIE (1.73a)

$$\eta \mathcal{T} \mathbf{J}_s + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{M}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}}. \quad (3.3)$$

The current \mathbf{J}_s and \mathbf{M}_s are assumed to verify the Impedance Boundary Condition

$$\mathbf{M}_s = -z^{\text{imp}} \mathcal{N} \mathbf{J}_s \quad (3.4)$$

where \mathcal{N} is an operator that can be either $\hat{\mathbf{n}} \times \mathcal{I}$ for the Leontovich IBC or \mathcal{N}_δ that is introduced later in this chapter. Substituting (3.4) into (3.3) results in the IBC-EFIE

$$\eta \mathcal{T} \mathbf{J}_s - z^{\text{imp}} \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathcal{N} \mathbf{J}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}}. \quad (3.5)$$

The IBC-EFIE operator is noted \mathcal{S} and defined as

$$\mathcal{S} = \eta \mathcal{T} - z^{\text{imp}} \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathcal{N}. \quad (3.6)$$

Then, the IBC-EFIE reads

$$\mathcal{S} \mathbf{J}_s = -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}}. \quad (3.7)$$

3.2.2 Discretization of the IBC-EFIE

The boundary Γ is discretized by a triangular mesh on which N_s RWG basis functions $(\mathbf{f}_n)_{n=1}^{N_s}$ (1.90) and N_s BC basis functions $(\mathbf{g}_n)_{n=1}^{N_s}$ (1.92) are constructed.

As it is done in the previous chapter, a mixed discretization is used for the currents. So, the surface electric current density \mathbf{J}_s is discretized by RWG basis functions \mathbf{f}_n and the surface magnetic current density \mathbf{M}_s is discretized by BC basis functions \mathbf{g}_n . The coefficients of the expansion are \mathbf{J} and $\mathbf{M} \in \mathbb{C}^{N_s}$ such that \mathbf{J}_s and \mathbf{M}_s are approximated by

$$\mathbf{J}_s \approx \sum_{n=1}^{N_s} [\mathbf{J}]_n \mathbf{f}_n \quad (3.8a)$$

$$\mathbf{M}_s \approx \sum_{n=1}^{N_s} [\mathbf{M}]_n \mathbf{g}_n. \quad (3.8b)$$

The EFIE (3.3) and the IBC (3.4) are tested with rotated RWG basis functions $\hat{\mathbf{n}} \times \mathbf{f}_n$ which effectively results in a linear system of $2N_s$ equations and $2N_s$ unknowns

$$\eta \mathbf{T} \mathbf{J} + \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K} \right) \mathbf{M} = \mathbf{E} \quad (3.9a)$$

$$\mathbf{G}_m \mathbf{M} = -z^{\text{imp}} \mathbf{N} \mathbf{J} \quad (3.9b)$$

where

$$[\mathbf{T}]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T} \mathbf{f}_n \rangle \quad (3.10a)$$

$$[\mathbf{K}]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{K} \mathbf{g}_n \rangle \quad (3.10b)$$

$$[\mathbf{G}_m]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathbf{g}_n \rangle \quad (3.10c)$$

$$[\mathbf{N}]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{N} \mathbf{f}_n \rangle \quad (3.10d)$$

$$[\mathbf{E}]_m = -\langle \mathbf{f}_m, \mathbf{E}^{\text{inc}} \rangle. \quad (3.10e)$$

3. Multiplicative preconditioner for a new IBC-EFIE

Note that if $\mathcal{N} = \hat{\mathbf{n}} \times \mathcal{I}$, then \mathbf{N} is equal to the Gram matrix \mathbf{G} as $\langle \hat{\mathbf{n}} \times \mathbf{f}_m, \hat{\mathbf{n}} \times \mathbf{f}_n \rangle = \langle \mathbf{f}_m, \mathbf{f}_n \rangle$. Again, the IBC-EFIE is solved in practice by substituting \mathbf{M} in (3.9)

$$\mathbf{S}\mathbf{J} = \mathbf{E} \quad (3.11a)$$

$$\mathbf{M} = -z^{\text{imp}} \mathbf{G}_m^{-1} \mathbf{N}\mathbf{J} \quad (3.11b)$$

where

$$\mathbf{S} = \eta \mathbf{T} - z^{\text{imp}} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K} \right) \mathbf{G}_m^{-1} \mathbf{N}. \quad (3.12)$$

The main observation here is that it is possible to modify the spectral properties of the IBC-EFIE by properly choosing the operator \mathcal{N} in the IBC, such that the solution of the IBC problem does not change significantly but the system is easier to precondition overall.

3.3 New Impedance Boundary Condition

3.3.1 Definition of the new IBC

The following IBC linking the surface electric and magnetic currents densities is assumed to hold on Γ

$$\mathbf{M}_s(\mathbf{r}) = -z^{\text{imp}} \mathcal{N}_\delta \mathbf{J}_s(\mathbf{r}) \quad (3.13)$$

where z^{imp} is a scalar complex impedance and where \mathcal{N}_δ is defined as

$$\mathcal{N}_\delta \mathbf{J}_s(\mathbf{r}) = \frac{2(1+i)}{\delta} \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-\frac{1+i}{\delta} |\mathbf{r}-\mathbf{r}'|}}{4\pi |\mathbf{r}-\mathbf{r}'|} \mathbf{J}_s(\mathbf{r}') dS'. \quad (3.14)$$

\mathcal{N}_δ is equal to \mathcal{T}_s (1.68b) rescaled and evaluated at a complex wave number such that $ik = (1+i)/\delta$,

$$\mathcal{N}_\delta = 2ik \mathcal{T}_s|_{ik=\frac{1+i}{\delta}}. \quad (3.15)$$

The function

$$\mathbf{r} \mapsto \frac{2(1+i)}{\delta} \frac{e^{-\frac{1+i}{\delta} |\mathbf{r}|}}{4\pi |\mathbf{r}|} \quad (3.16)$$

is a nascent Dirac distribution on Γ when $\delta \rightarrow 0$. In other words,

$$\mathcal{N}_\delta \rightarrow \hat{\mathbf{n}} \times \mathcal{I} \text{ when } \delta \rightarrow 0. \quad (3.17)$$

So, if δ is sufficiently small, the new IBC (3.13) is in fact an approximation of the Leontovich IBC (3.1).

3.3.2 Validity of the new IBC

The starting point to derive the new IBC is the EFIE in the interior domain (1.75a)

$$-\eta' \mathcal{T}' \mathbf{J}_s + \left(\frac{\mathcal{I}}{2} + \mathcal{K}' \right) \mathbf{M}_s = \mathbf{0}. \quad (3.18)$$

The interior medium is characterized by a permittivity ε' , a permeability μ' , an impedance $\eta' = \sqrt{\mu'/\varepsilon'}$ and a wave number $k' = \omega\sqrt{\mu'\varepsilon'}$. Assume that the interior medium has a high conductivity σ' and that the frequency is low enough to neglect the real part of the permittivity

$$\varepsilon' = \varepsilon_0 + \frac{\sigma'}{i\omega} \approx \frac{\sigma'}{i\omega}. \quad (3.19)$$

Also, Γ is assumed to be smooth. The frequency should be high enough such that the skin depth

$$\delta_{sd} = \sqrt{\frac{2}{\omega\mu'\sigma'}} \quad (3.20)$$

is much smaller than the radius of curvature of Γ . In this regime, the surface integral operators verify

$$\frac{\mathcal{I}}{2} + \mathcal{K}' \approx \frac{\mathcal{I}}{2} \quad (3.21a)$$

$$\mathcal{T}' \approx -ik' \mathcal{T}'_s \Big|_{ik' = \frac{1+i}{\delta_{sd}}}. \quad (3.21b)$$

The first approximation (3.21a) is valid because Γ is assumed to be smooth. Actually, the equality holds only if Γ is perfectly planar i.e. an infinite plane for which $\mathcal{K}' = 0$. The second approximation (3.21b) uses the assumption (3.19) to rewrite the wave number as a function of the skin depth

$$ik' = i\omega\sqrt{\mu'\varepsilon'} \approx \frac{1+i}{\delta_{sd}}. \quad (3.22)$$

If the skin depth is small enough then $ik' \mathcal{T}'_s$ becomes dominant compared to $\frac{1}{ik'} \mathcal{T}'_h$ in (1.68a) so that the hypersingular part is neglected in (3.21b). Then, using (3.21) in (3.18) results in the new boundary condition

$$\mathbf{M}_s = \eta' \left(\frac{\mathcal{I}}{2} + \mathcal{K}' \right)^{-1} \mathcal{T}' \mathbf{J}_s \quad (3.23a)$$

$$\approx -\eta' \mathcal{N}_{\delta_{sd}} \mathbf{J}_s. \quad (3.23b)$$

To compare with the Leontovich IBC, $\mathbf{M}_s = -z^{\text{imp}} \hat{\mathbf{n}} \times \mathbf{J}_s$ corresponds to the stronger assumption that $\delta_{sd} = 0$. Note that in practice, in the new operator \mathcal{N}_δ , δ is not always chosen to be equal to the skin depth δ_{sd} because

- the smaller δ is, the more oscillating is the integration kernel of \mathcal{N}_δ , so δ should be in the order of the average edge length h or larger to compute \mathbf{N} accurately,

- δ should be large enough to benefit from the spectral properties of \mathcal{N}_δ in the preconditioning,
- the scale of the variations in the solution is not in order of the skin depth but in the order of magnitude of the wavelength or the radius of curvature of the geometry.

3.3.3 Accuracy of the new IBC

The definition of the new operator \mathcal{N}_δ results from an approximation of the surface integral operators \mathcal{K}' and \mathcal{T}' in the interior domain. In this chapter, the definition

$$\mathcal{N}_\delta = 2ik\mathcal{T}_s|_{ik=\frac{1+i}{\delta}} \quad (3.24)$$

is used, but two other forms for \mathcal{N}_δ have also been considered. The first one is

$$\mathcal{N}_\delta = -2\mathcal{T}|_{ik=\frac{1+i}{\delta}} \quad (3.25)$$

that keeps the entire operator \mathcal{T} . Regarding the derivation of \mathcal{N}_δ in the previous section, it is the form that does as few approximations as possible. For example, with this form in the IBC-EFIE it is possible to approximate a PMCHWT formulation where the scatterer is highly conducting ($\epsilon' \approx \sigma'/(i\omega)$) and for which the operator \mathcal{K}' is neglected in the interior domain. The other form considered is

$$\mathcal{N}_\delta = 2ik\mathcal{T}_s|_{ik=\frac{1}{\delta}} \quad (3.26)$$

which is a cruder approximation as it keeps only the real part of $\frac{1+i}{\delta}$ as a wave number, but has the computational benefit of being purely real. Note that these two alternatives for \mathcal{N}_δ are also valid regarding the preconditioning of the IBC-EFIE because they both have the spectral properties of the operator \mathcal{T}_s that are leveraged in this chapter. Practically, the new operator \mathcal{N}_δ is based on the operator \mathcal{T} so it is easy to compute from an existing code.

Like the Leontovich IBC, this new IBC is valid on smooth boundaries on which the solution varies at a much larger scale than the skin depth. Regarding the accuracy in presence of sharp edges, wedges or corners, it is known that the classic $\mathbf{M}_s = -z^{\text{imp}}\hat{\mathbf{n}} \times \mathbf{J}_s$ is not an accurate model [14–16], so the new IBC is not expected to behave well in these cases either.

3.4 Vector Spherical Harmonics analysis

In this section, the VSH \mathbf{X}_{lm} and \mathbf{U}_{lm} (defined in section 1.5.2.1) are used to study the new IBC and the properties of the IBC-EFIE. The surface Γ is a sphere of radius a . First, the new boundary operator \mathcal{N}_δ is expanded in terms of VSH. Then, an analytic solution can be constructed for the new IBC-EFIE. Finally, the asymptotic behaviour of the IBC-EFIE is studied at low frequency and for dense meshes.

3.4.1 Vector Spherical Harmonics expansion of the new IBC operator

The VSH expansion of \mathcal{N}_δ can be obtained from the VSH expansion of \mathcal{T}_s thanks to (3.15). While the expansion of \mathcal{T} is well known and readily available in the literature because it is the PEC-EFIE operator [28], the studies on \mathcal{T}_s or \mathcal{T}_h are scarce. Therefore, in the following $\mathcal{T}_h \mathbf{X}_{lm}$ and $\mathcal{T}_h \mathbf{U}_{lm}$ are first computed. Then, once the expansion of \mathcal{T}_h is known, $\mathcal{T}_s \mathbf{X}_{lm}$ and $\mathcal{T}_s \mathbf{U}_{lm}$ are computed using (1.68a)

$$ik\mathcal{T}_s = -\mathcal{T} + \frac{1}{ik}\mathcal{T}_h \quad (3.27)$$

and the expansion of \mathcal{T} (1.131)

$$\mathcal{T} \mathbf{X}_{lm} = -J_l(ka)H_l^{(2)}(ka)\mathbf{U}_{lm} \quad (3.28a)$$

$$\mathcal{T} \mathbf{U}_{lm} = J_l'(ka)H_l^{(2)'}(ka)\mathbf{X}_{lm}. \quad (3.28b)$$

\mathbf{X}_{lm} is solenoidal so from $\nabla \cdot \mathbf{X}_{lm} = 0$ it follows that $\mathcal{T}_h \mathbf{X}_{lm} = \mathbf{0}$ and

$$ik\mathcal{T}_s \mathbf{X}_{lm} = -\mathcal{T} \mathbf{X}_{lm} = J_l(ka)H_l^{(2)}(ka)\mathbf{U}_{lm}. \quad (3.29)$$

The computation of $\mathcal{T}_h \mathbf{U}_{lm}$ has some prerequisites related to spherical harmonics. First, the spherical harmonics are eigenfunctions of the scalar laplacian on the sphere

$$\nabla \cdot \nabla Y_{lm} = \nabla^2 Y_{lm} = -\frac{l(l+1)}{a^2} Y_{lm}. \quad (3.30)$$

And second, the spherical harmonics expansion of the Green's function is given for all $\mathbf{r}, \mathbf{r}' \in \mathbb{R}^3$ such that $|\mathbf{r}'| < |\mathbf{r}|$ by [41]

$$\frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} = -ik \sum_{l=0}^{+\infty} j_l(k|\mathbf{r}'|) h_l^{(2)}(k|\mathbf{r}|) \sum_{m=-l}^l \overline{Y_{lm}(\hat{\mathbf{r}}')} Y_{lm}(\hat{\mathbf{r}}) \quad (3.31)$$

where j_l are the spherical Bessel functions and $h_l^{(2)}$ are the spherical Hankel functions of second kind. Then, using the orthonormality of the spherical harmonics

$$\iint_{\mathbf{r}' \in \Gamma} Y_{lm}(\hat{\mathbf{r}}') \overline{Y_{l'm'}(\hat{\mathbf{r}}')} \frac{dS'}{a^2} = \delta_{l,l'} \delta_{m,m'} \quad (3.32)$$

results in

$$\iint_{\mathbf{r}' \in \Gamma} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} Y_{lm}(\hat{\mathbf{r}}') \frac{dS'}{a^2} = -ik j_l(ka) h_l^{(2)}(k|\mathbf{r}|) Y_{lm}(\hat{\mathbf{r}}). \quad (3.33)$$

The limit of the above expression when \mathbf{r} approaches Γ (i.e. when $|\mathbf{r}| \rightarrow a$ with $|\mathbf{r}| > a$) is

$$\iint_{\mathbf{r}' \in \Gamma} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} Y_{lm}(\hat{\mathbf{r}}') \frac{dS'}{a^2} = -ik j_l(ka) h_l^{(2)}(ka) Y_{lm}(\hat{\mathbf{r}}). \quad (3.34)$$

With (3.30) and (3.34), the computation of $\mathcal{T}_h \mathbf{U}_{lm}$ is pretty straightforward

$$\mathcal{T}_h \mathbf{U}_{lm} = \mathcal{T}_h \left(-\frac{a}{i\sqrt{l(l+1)}} \nabla Y_{lm} \right) \quad (3.35a)$$

$$= -\frac{a}{i\sqrt{l(l+1)}} \hat{\mathbf{n}} \times \nabla \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \nabla \cdot \nabla Y_{lm}(\hat{\mathbf{r}}') dS' \quad (3.35b)$$

$$= \frac{a}{i\sqrt{l(l+1)}} l(l+1) \hat{\mathbf{n}} \times \nabla \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} Y_{lm}(\hat{\mathbf{r}}') \frac{dS'}{a^2} \quad (3.35c)$$

$$= -\frac{a}{i\sqrt{l(l+1)}} ikl(l+1) j_l(ka) h_l^{(2)}(ka) \hat{\mathbf{n}} \times \nabla Y_{lm}(\hat{\mathbf{r}}) \quad (3.35d)$$

$$= -ikl(l+1) j_l(ka) h_l^{(2)}(ka) \mathbf{X}_{lm}. \quad (3.35e)$$

This results in the expression for $\mathcal{T}_s \mathbf{U}_{lm}$

$$ik\mathcal{T}_s \mathbf{U}_{lm} = -\mathcal{T} \mathbf{U}_{lm} + \frac{1}{ik} \mathcal{T}_h \mathbf{U}_{lm} \quad (3.36a)$$

$$= -\left(J_l'(ka) H_l^{(2)'}(ka) + l(l+1) j_l(ka) h_l^{(2)}(ka) \right) \mathbf{X}_{lm}. \quad (3.36b)$$

In fact, these results are also valid for a complex wave number such that the VSH expansion of \mathcal{N}_δ is simply obtained with the substitution $ik = (1+i)/\delta$ as in (3.15)

$$\langle \mathbf{U}_{lm}, \mathcal{N}_\delta \mathbf{X}_{lm} \rangle = 2J_l(u) H_l^{(2)}(u) \quad (3.37a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{N}_\delta \mathbf{U}_{lm} \rangle = -2 \left(J_l'(u) H_l^{(2)'}(u) + l(l+1) j_l(u) h_l^{(2)}(u) \right) \quad (3.37b)$$

where

$$u = (1-i) \frac{a}{\delta}. \quad (3.38)$$

3.4.2 Analytic solution for the currents

Similarly to the previous chapter (see section 2.3.2), an analytic solution in terms of VSH is built for the currents. Using (3.37), (1.131) and (1.132), the VSH expansion of the operator \mathcal{S} (3.6) is

$$\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle = -\left(\eta H_l^{(2)}(ka) + z^{\text{imp}} i H_l^{(2)'}(ka) \langle \mathbf{U}_{lm}, \mathcal{N}_\delta \mathbf{X}_{lm} \rangle \right) J_l(ka) \quad (3.39a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{S} \mathbf{U}_{lm} \rangle = \left(\eta H_l^{(2)'}(ka) + z^{\text{imp}} i H_l^{(2)}(ka) \langle \mathbf{X}_{lm}, \mathcal{N}_\delta \mathbf{U}_{lm} \rangle \right) J_l'(ka). \quad (3.39b)$$

Similarly to the assumption done in the previous chapter analysis, the impedance z^{imp} and the frequency f are assumed to be different than the resonant impedances or frequencies that cancel $\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle$ and $\langle \mathbf{X}_{lm}, \mathcal{S} \mathbf{U}_{lm} \rangle$. The currents are expanded in VSH as

$$\mathbf{J}_s = \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle \mathbf{X}_{lm} + \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle \mathbf{U}_{lm} \quad (3.40a)$$

$$\mathbf{M}_s = \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{X}_{lm}, \mathbf{M}_s \rangle \mathbf{X}_{lm} + \sum_{l=1}^{+\infty} \sum_{m=-l}^l \langle \mathbf{U}_{lm}, \mathbf{M}_s \rangle \mathbf{U}_{lm} \quad (3.40b)$$

where the coefficients are given by (see (2.21))

$$\langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle = \frac{\langle \mathbf{U}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle}{\langle \mathbf{U}_{lm}, \mathcal{S}\mathbf{X}_{lm} \rangle} \quad (3.41a)$$

$$\langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle = \frac{\langle \mathbf{X}_{lm}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle}{\langle \mathbf{X}_{lm}, \mathcal{S}\mathbf{U}_{lm} \rangle} \quad (3.41b)$$

$$\langle \mathbf{X}_{lm}, \mathbf{M}_s \rangle = -z^{\text{imp}} \langle \mathbf{X}_{lm}, \mathcal{N}_\delta \mathbf{U}_{lm} \rangle \langle \mathbf{U}_{lm}, \mathbf{J}_s \rangle \quad (3.41c)$$

$$\langle \mathbf{U}_{lm}, \mathbf{M}_s \rangle = -z^{\text{imp}} \langle \mathbf{U}_{lm}, \mathcal{N}_\delta \mathbf{X}_{lm} \rangle \langle \mathbf{X}_{lm}, \mathbf{J}_s \rangle. \quad (3.41d)$$

The values for the coefficients of the excitation have been computed in the previous chapter (see (2.25)) for the plane wave $\mathbf{E}^{\text{inc}}(\mathbf{r}) = E_0 e^{ikz} \hat{\mathbf{x}}$

$$\langle \mathbf{X}_{l,\pm 1}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle = \pm 2\pi E_0 i^l \sqrt{\frac{2l+1}{4\pi}} \frac{J'_l(ka)}{ka} \quad (3.42a)$$

$$\langle \mathbf{U}_{l,\pm 1}, -\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}} \rangle = -2\pi E_0 i^l \sqrt{\frac{2l+1}{4\pi}} j_l(ka). \quad (3.42b)$$

3.4.3 Low frequency behaviour

Regarding the physics of the skin effect, if δ is set to be equal to the skin depth, then it should also vary with the frequency. Thus the coefficients $\langle \mathbf{U}_{lm}, \mathcal{N}_\delta \mathbf{X}_{lm} \rangle$ and $\langle \mathbf{X}_{lm}, \mathcal{N}_\delta \mathbf{U}_{lm} \rangle$ should also vary with the frequency. But this is true as long as δ remains much smaller than the radius of curvature of the scatterer (the radius a in the case of a sphere). So it can be assumed that $a/\delta \gg 1$. Regarding (3.17), it is not surprising that the limits of (3.37) when $a/\delta \rightarrow +\infty$ are $\langle \mathbf{U}_{lm}, \hat{\mathbf{n}} \times \mathbf{X}_{lm} \rangle = 1$ and $\langle \mathbf{X}_{lm}, \hat{\mathbf{n}} \times \mathbf{U}_{lm} \rangle = -1$. Therefore, in the following low frequency study, it is assumed that

$$\langle \mathbf{U}_{lm}, \mathcal{N}_\delta \mathbf{X}_{lm} \rangle \approx 1 \quad (3.43a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{N}_\delta \mathbf{U}_{lm} \rangle \approx -1 \quad (3.43b)$$

as one would obtain with the classic $\hat{\mathbf{n}} \times$ in the Leontovich IBC. In other words, the new IBC does not modify the low frequency behaviour compared to the Leontovich IBC. In fact, the approximations in (3.43) only hold for small orders l whereas for higher l these coefficients vary proportionally to $1/l$ as it is explained in the next section.

The impedance z^{imp} is assumed to be far from the resonant impedances that cancel $\langle \mathbf{U}_{lm}, \mathcal{S}\mathbf{X}_{lm} \rangle$ and $\langle \mathbf{X}_{lm}, \mathcal{S}\mathbf{U}_{lm} \rangle$ in (3.39). Then, using the asymptotic behaviour of the Bessel family functions for small arguments, the low frequency scaling of the IBC-EFIE operator is

$$\langle \mathbf{U}_{lm}, \mathcal{S}\mathbf{X}_{lm} \rangle \underset{k \rightarrow 0}{\sim} - \left(\frac{ika}{l} \eta + z^{\text{imp}} \right) \frac{l}{2l+1} \quad (3.44a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{S}\mathbf{U}_{lm} \rangle \underset{k \rightarrow 0}{\sim} \left(\frac{l}{ika} \eta - z^{\text{imp}} \right) \frac{l+1}{2l+1}. \quad (3.44b)$$

In all practical models for the impedance, z^{imp} is negligible compared to $\eta/(ka)$, so in (3.44b), z^{imp} can be neglected as is done in the previous chapter.

3.4.4 Dense mesh behaviour

Using the asymptotic behaviour of the Bessel family functions for large order, the new operator \mathcal{N}_δ (3.37) scales as

$$\langle \mathbf{U}_{lm}, \mathcal{N}_\delta \mathbf{X}_{lm} \rangle \underset{l \rightarrow +\infty}{\sim} \frac{i u}{l} \quad (3.45a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{N}_\delta \mathbf{U}_{lm} \rangle \underset{l \rightarrow +\infty}{\sim} -\frac{i u}{l}. \quad (3.45b)$$

This $1/l$ behaviour of \mathcal{N}_δ is in fact inherited from the operator \mathcal{T}_s . In particular, it can be computed from (3.29) that

$$\langle \mathbf{U}_{lm}, \mathcal{T}_s \mathbf{X}_{lm} \rangle \underset{l \rightarrow +\infty}{\sim} \frac{a}{2l}. \quad (3.46)$$

Also, note that from (3.35e),

$$\langle \mathbf{X}_{lm}, \mathcal{T}_h \mathbf{U}_{lm} \rangle \underset{l \rightarrow +\infty}{\sim} \frac{l}{2a}. \quad (3.47)$$

Before computing the asymptotic behaviour of the IBC-EFIE operator \mathcal{S} , it should be noted that the resonant impedances associated to $\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle$ in (3.39) tend to a finite limit when $l \rightarrow +\infty$. In fact, $\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle = 0$ if

$$z^{\text{imp}} = -\frac{\eta H_l^{(2)}(ka)}{i H_l^{(2)'}(ka) \langle \mathbf{U}_{lm}, \mathcal{N}_\delta \mathbf{X}_{lm} \rangle} \quad (3.48a)$$

$$\underset{l \rightarrow +\infty}{\rightarrow} -\frac{ka}{u} \eta = -\frac{1+i}{2} k \delta \eta. \quad (3.48b)$$

So in the following asymptotic behaviour of the IBC-EFIE operator \mathcal{S} , the impedance $z^{\text{imp}} = -\frac{ka}{u} \eta$ is excluded

$$\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle \underset{l \rightarrow +\infty}{\sim} -\frac{1}{2} \left(\frac{i k a}{l} \eta + \frac{i u}{l} z^{\text{imp}} \right) \quad (3.49a)$$

$$\langle \mathbf{X}_{lm}, \mathcal{S} \mathbf{U}_{lm} \rangle \underset{l \rightarrow +\infty}{\sim} \frac{1}{2} \left(\frac{l}{i k a} \eta - \frac{i u}{l} z^{\text{imp}} \right) \underset{l \rightarrow +\infty}{\sim} \frac{l}{2 i k a} \eta. \quad (3.49b)$$

Note that $\langle \mathbf{U}_{lm}, \mathcal{S} \mathbf{X}_{lm} \rangle$ scales proportionally to $1/l$, contrary to the Leontovich IBC-EFIE where it tends to a constant when $z^{\text{imp}} \neq 0$ (see (2.34a)). So it is clear that this part of the spectrum (non-solenoidal) can always be preconditioned with a Calderón preconditioning by applying the operator \mathcal{T}_h that scales proportionally to l (see (3.47)). This is opposed to the Leontovich IBC-EFIE that does not need a regularization on the non-solenoidal part of the spectrum (see its dense mesh preconditioner (2.46)).

Regarding the solenoidal part of the spectrum, it can be regularized by applying on it the operator \mathcal{T}_s that scales proportionally to $1/l$ (see (3.46)). This operator \mathcal{T}_s also applies on the non-solenoidal part for an overall $1/l^2$ scaling, but it is dominated by the constant scaling due to the Calderón preconditioning with \mathcal{T}_h on this part, as explained above.

3.5 Construction of a Calderón multiplicative preconditioner

Based on the analysis in the previous section, a Calderón multiplicative preconditioner can be constructed for the new IBC-EFIE to solve both the low frequency breakdown and the dense mesh breakdown. Contrary to the previous chapter where the quasi-Helmholtz projectors are used to precondition the Leontovich IBC-EFIE, they are not used here. Consequently, at very low frequency the numerical cancellations occur in this new formulation. However, in the frequency regime where the IBC holds, this should manifest as a lost in digits of precision so it does not prevent from retrieving the solution.

In the previous section, it is explained that the preconditioner should contain the operator \mathcal{T}_s to act on the solenoidal part of the spectrum ($\langle \mathbf{X}_{lm}, \mathbf{S}\mathbf{U}_{lm} \rangle$), and the operator \mathcal{T}_h to act on the non-solenoidal part of the spectrum ($\langle \mathbf{U}_{lm}, \mathbf{S}\mathbf{X}_{lm} \rangle$). Therefore, a preconditioner that is a linear combination of \mathcal{T}_s and \mathcal{T}_h is investigated

$$\tilde{\mathcal{T}} = \alpha \mathcal{T}_s + \beta \mathcal{T}_h \quad (3.50)$$

where α and β are unknown coefficients that are computed in the following. Applying \mathcal{T}_s (3.46) on $\mathbf{S}\mathbf{U}_{lm}$ (3.49b) results in the following scaling for dense mesh and low frequency

$$\langle \mathbf{U}_{lm}, \mathcal{T}_s \mathbf{S}\mathbf{U}_{lm} \rangle \underset{k \rightarrow 0}{\overset{l \rightarrow +\infty}{\sim}} \frac{a}{2l} \frac{l}{2ika} \quad (3.51a)$$

$$\underset{k \rightarrow 0}{\overset{l \rightarrow +\infty}{\sim}} \left(\frac{1}{-ik} \right) \left(\frac{-\eta}{4} \right). \quad (3.51b)$$

To compensate the $1/(-ik)$ scaling, $\alpha = -ik$ is chosen. In fact, it is the same scaling as the Calderón multiplicative preconditioner of the PEC-EFIE [7]. The $-\eta/4$ scaling that was taken away in the choice of α should be factorized away also in the determination of β . Applying the operator \mathcal{T}_h (3.47) on $\mathbf{S}\mathbf{X}_{lm}$ (3.49a) results in

$$\langle \mathbf{X}_{lm}, \mathcal{T}_h \mathbf{S}\mathbf{X}_{lm} \rangle \underset{k \rightarrow 0}{\overset{l \rightarrow +\infty}{\sim}} -\frac{l}{2a} \frac{1}{2} \left(\frac{ika}{l} \eta + \frac{i u}{l} z^{\text{imp}} \right) \quad (3.52a)$$

$$\underset{k \rightarrow 0}{\overset{l \rightarrow +\infty}{\sim}} \left(ik + \frac{i u}{a} \frac{z^{\text{imp}}}{\eta} \right) \left(\frac{-\eta}{4} \right) \quad (3.52b)$$

$$\underset{k \rightarrow 0}{\overset{l \rightarrow +\infty}{\sim}} \left(ik + \frac{1+i}{\delta} \frac{z^{\text{imp}}}{\eta} \right) \left(\frac{-\eta}{4} \right). \quad (3.52c)$$

Thus $\beta = (ik + (1+i)z^{\text{imp}}/(\delta\eta))^{-1}$ is chosen. To sum up, the operator that preconditions the new IBC-EFIE is

$$\tilde{\mathcal{T}} = -ik \mathcal{T}_s + \frac{1}{ik + \frac{1+i}{\delta} \frac{z^{\text{imp}}}{\eta}} \mathcal{T}_h. \quad (3.53)$$

3. Multiplicative preconditioner for a new IBC-EFIE

In fact, this is the operator \mathcal{T} but with a different scaling on the hypersingular part. In addition, it is clear that $\tilde{\mathcal{T}} \rightarrow \mathcal{T}$ when $z^{\text{imp}} \rightarrow 0$ i.e. the new preconditioned IBC-EFIE reduces to the standard Calderón preconditioned PEC-EFIE [7].

A conforming discretization of (3.53) is obtained using the BC basis functions \mathbf{g}_n ($n \in [1, N_s]$) and reads

$$\tilde{\mathbb{T}} = -ik\mathbb{T}_s + \frac{1}{ik + \frac{1+i}{\delta} \frac{z^{\text{imp}}}{\eta}} \mathbb{T}_h \quad (3.54)$$

where $[\mathbb{T}_s]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{g}_m, \mathcal{T}_s \mathbf{g}_n \rangle$ and $[\mathbb{T}_h]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{g}_m, \mathcal{T}_h \mathbf{g}_n \rangle$. As usual, the preconditioner (3.54) discretized with the BC basis functions \mathbf{g}_n is not directly multiplied to the IBC-EFIE system (3.11a) that is tested with rotated RWG basis functions $\hat{\mathbf{n}} \times \mathbf{f}_n$. Indeed, the inverse of the mix-Gram matrix \mathbf{G}_m ($[\mathbf{G}_m]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathbf{g}_n \rangle$) must be inserted between the two discretizations. Therefore, the preconditioned system is

$$\tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{S} \mathbf{J} = \tilde{\mathbb{T}} \mathbf{G}_m^{-1} \mathbf{E}. \quad (3.55)$$

As before, once the electric current coefficients \mathbf{J} have been found, the magnetic current coefficients \mathbf{M} are computed using

$$\mathbf{M} = -z^{\text{imp}} \mathbf{G}_m^{-1} \mathbf{N} \mathbf{J}. \quad (3.56)$$

3.6 Numerical results

The numerical results are divided in 3 parts. The first part is related to the new IBC and its accuracy. The second part is related to the stability of the new preconditioned formulation. And the third part shows some realistic scenarios.

3.6.1 Results related to the new IBC accuracy

In this section, the RCS are computed in the xz plane. The excitation is a plane wave that is polarized on $\hat{\mathbf{x}}$ and the direction of propagation is $\hat{\mathbf{k}} = -\hat{\mathbf{z}}$ as

$$\mathbf{E}^{\text{inc}}(\mathbf{r}) = E_0 e^{ik\hat{\mathbf{z}} \cdot \mathbf{r}} \hat{\mathbf{x}} \quad (3.57)$$

where the peak amplitude is $E_0 = 1$ V/m. It corresponds to a so-called VV-polarization.

The first example compares the RCS obtained with the classical Leontovich IBC $\mathbf{M}_s = -z^{\text{imp}} \hat{\mathbf{n}} \times \mathbf{J}_s$ (using the formulation in the chapter 2) and the new IBC $\mathbf{M}_s = -z^{\text{imp}} \mathcal{N}_\delta \mathbf{J}_s$. In the figure 3.1, the IBC-EFIEs and the analytic solutions for the two IBCs have been plotted for a unit sphere at $f = 100$ MHz with $h \approx \delta = 0.15$ m and $z^{\text{imp}} = (0.8 + 0.8i)\eta$. In this case, δ is sufficiently small for the two formulations to yield comparable results.

As it has been shown in (3.17), the new IBC operator \mathcal{N}_δ tends to $\hat{\mathbf{n}} \times \mathcal{I}$ (Leontovich IBC) when $\delta \rightarrow 0$. On the figure 3.2, the analytic RCSs for the new IBC have been plotted for decreasing δ ($f = 200$ MHz, $z^{\text{imp}} = (0.4 + 0.4i)\eta$). It can be seen that the solution obtained with the new IBC effectively converges to the classical

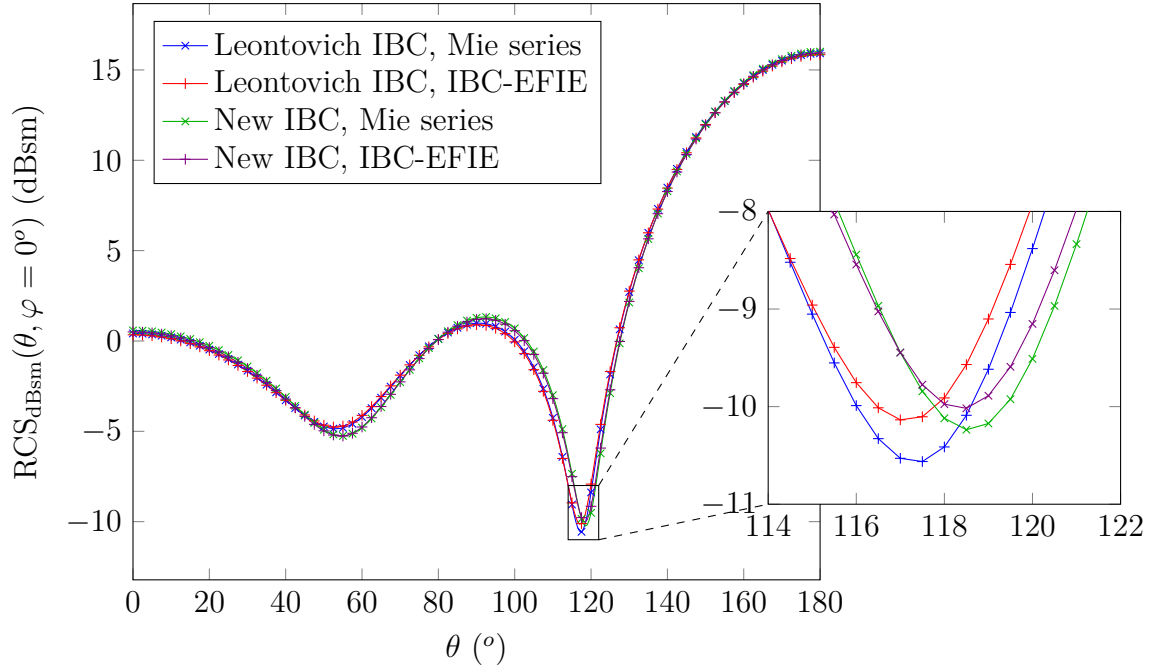


Figure 3.1: Comparison of the Leontovich IBC and the new IBC, numeric (IBC-EFIE) and analytic (Mie series) solutions (unit sphere, $f = 100$ MHz, $h \approx \delta = 0.15$ m, $z^{\text{imp}} = (0.8 + 0.8i)\eta$)

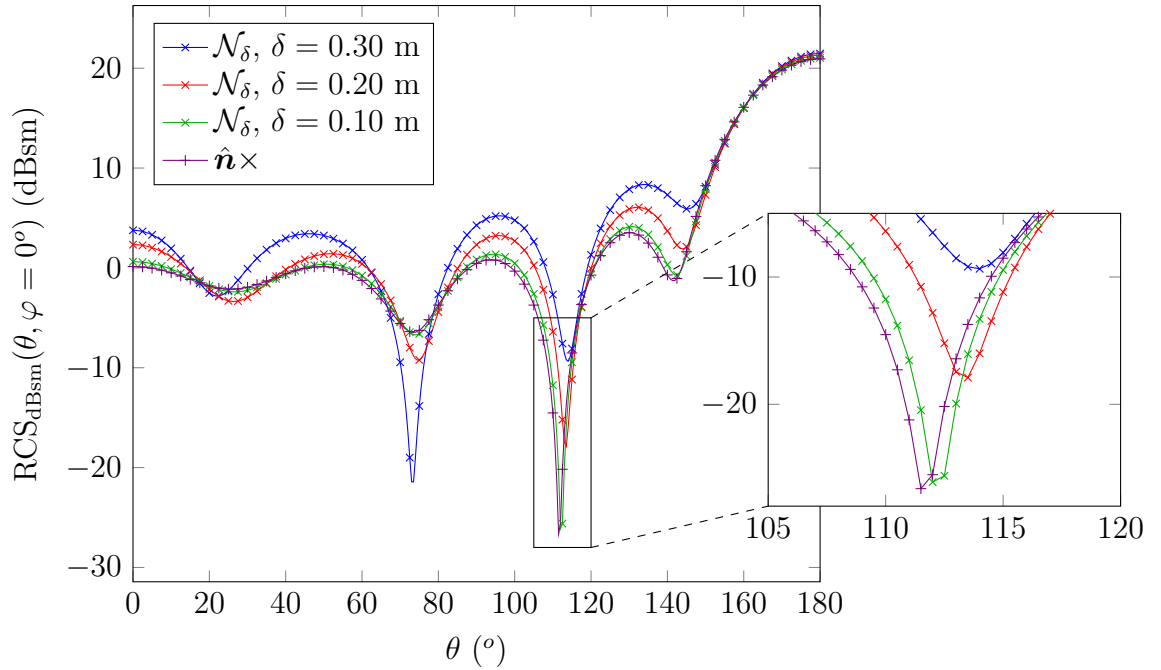


Figure 3.2: Analytic comparison of the RCS obtained for different IBCs (unit sphere, $f = 200$ MHz, $z^{\text{imp}} = (0.4 + 0.4i)\eta$)

Leontovich IBC. Note that these RCSs are analytic solutions. In practice, δ cannot be chosen arbitrary low because it makes the integration kernel of \mathcal{N}_δ very oscillatory (it is essentially \mathcal{T}_s with a wave number $ik' = (1 + i)/\delta$) and thus difficult to integrate accurately.

3.6.2 Results related to the stability of the formulation

In this section, the condition number and iteration count required to solve the systems are plotted as functions of the frequency f and the average edge length h . It is the number of iterations required to reach a residual of 10^{-8} with a GMRES solver (restarted every 20 iterations, limited to N_s iterations). There are 4 formulations on each graphs:

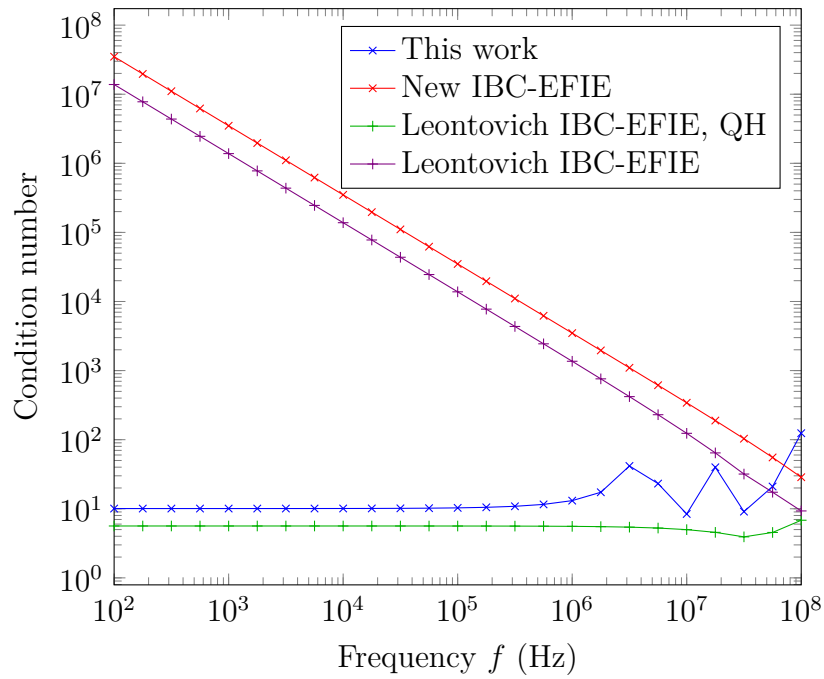
- "This work", is the formulation (3.55) developed in this chapter, which is the new IBC-EFIE that is preconditioned with a Calderón-like multiplicative preconditioner.
- "New IBC-EFIE" is the formulation (3.11a) developed in this chapter without the preconditioning.
- "Leontovich IBC-EFIE, QH", is the formulation (2.47) developed in the chapter 2, which is the low frequency and dense mesh stable Leontovich IBC-EFIE stabilized with quasi-Helmholtz projectors.
- "Leontovich IBC-EFIE", is the non-stabilized Leontovich IBC-EFIE (2.13).

The figure 3.3 shows the variations of the condition number and the number of iterations as functions of the frequency when the impedance is constant (unit sphere, $N_s = 1920$, $h \approx \delta = 0.15$ m, $z^{\text{imp}} = (0.8 + 0.8i)\eta$). It is clear from these graphs that the formulation is low frequency stable as the condition number and the number of iterations are relatively low. Note that the low frequency breakdown in the non-preconditioned formulation (3.11a) manifests as a growing of the condition number proportional to the inverse of the frequency, which is similar to the Leontovich IBC-EFIE (2.13).

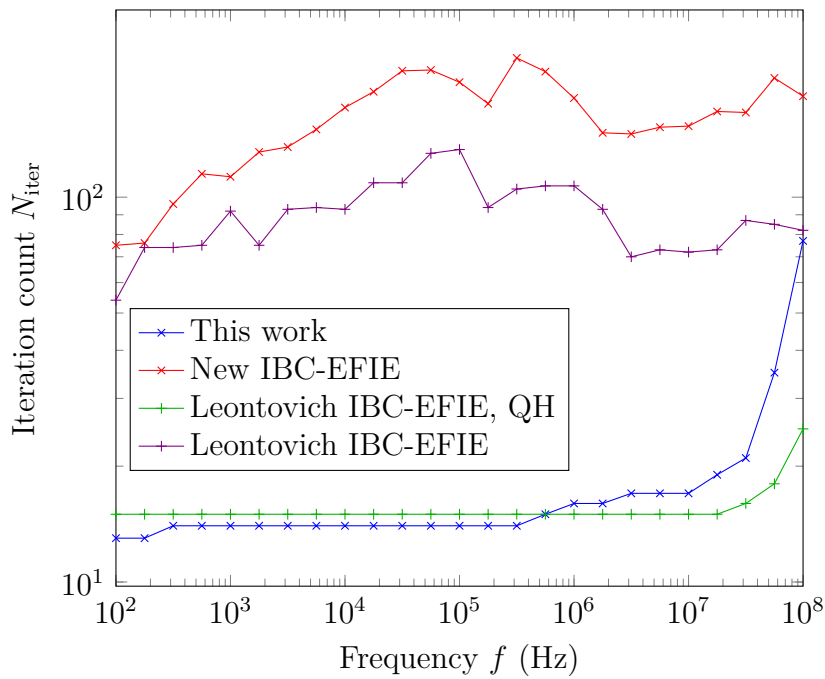
Similarly, the figure 3.4 shows the variations of the condition number and the number of iterations as functions of the frequency but the impedance is varying with the frequency (unit sphere, $N_s = 1920$, $h \approx \delta = 0.15$ m). In this figure 3.4, the impedance is proportional to the square root of the frequency f such that

$$z^{\text{imp}} = \sqrt{\frac{\mu_0 \omega}{2\sigma'}}(1 + i) \quad (3.58)$$

where the angular frequency is $\omega = 2\pi f$ and the conductivity is set to $\sigma' = 1$ S/m. Note that in this case, since $|z^{\text{imp}}| \gg \eta ka$, the condition numbers of the non-preconditioned formulations grow proportionally to $1/(|z^{\text{imp}}| k) \propto k^{-1.5}$ according to (2.30).



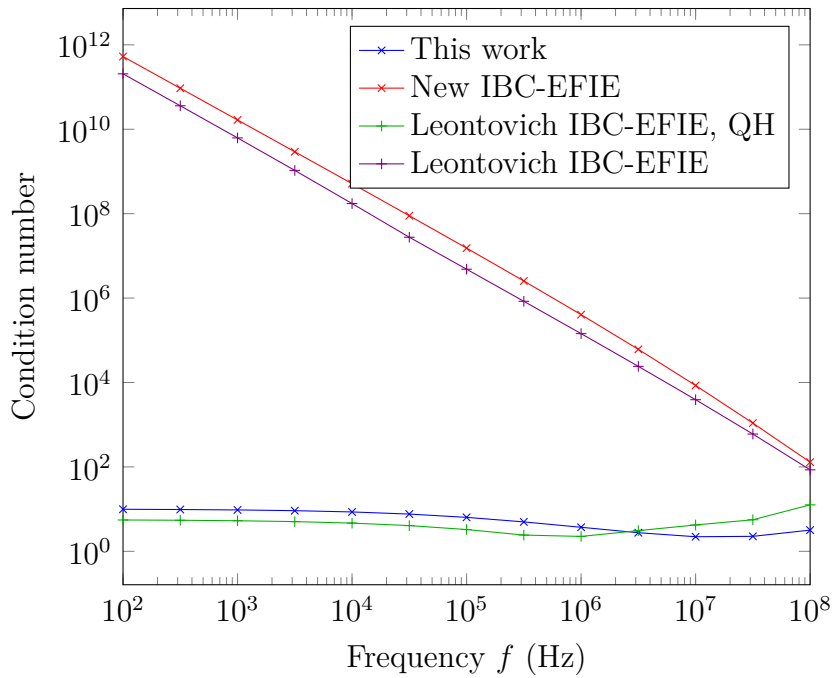
(a) Condition number as a function of the frequency



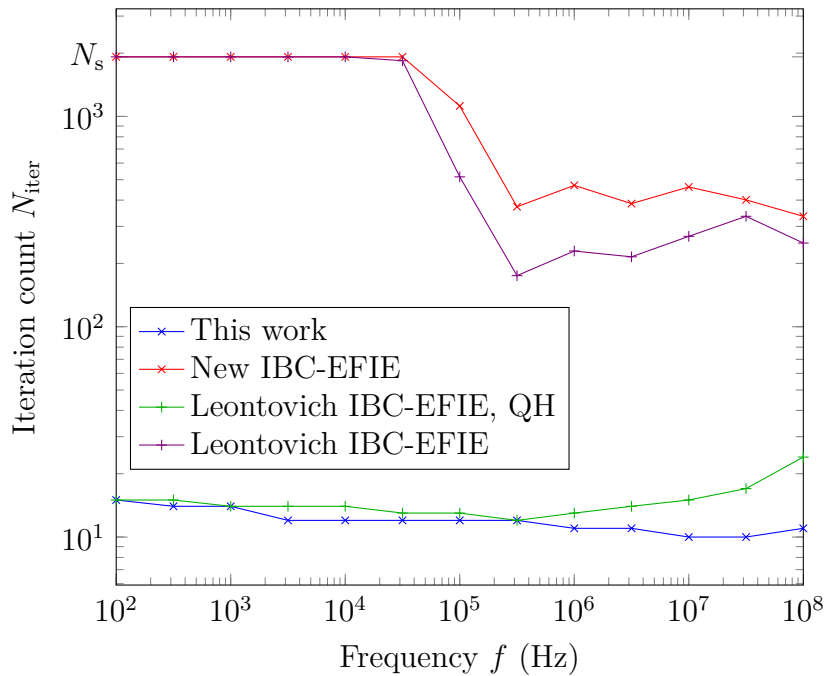
(b) Iteration count as a function of the frequency

Figure 3.3: Condition number and iteration count as functions of the frequency f when the impedance z^{imp} is constant

3. Multiplicative preconditioner for a new IBC-EFIE

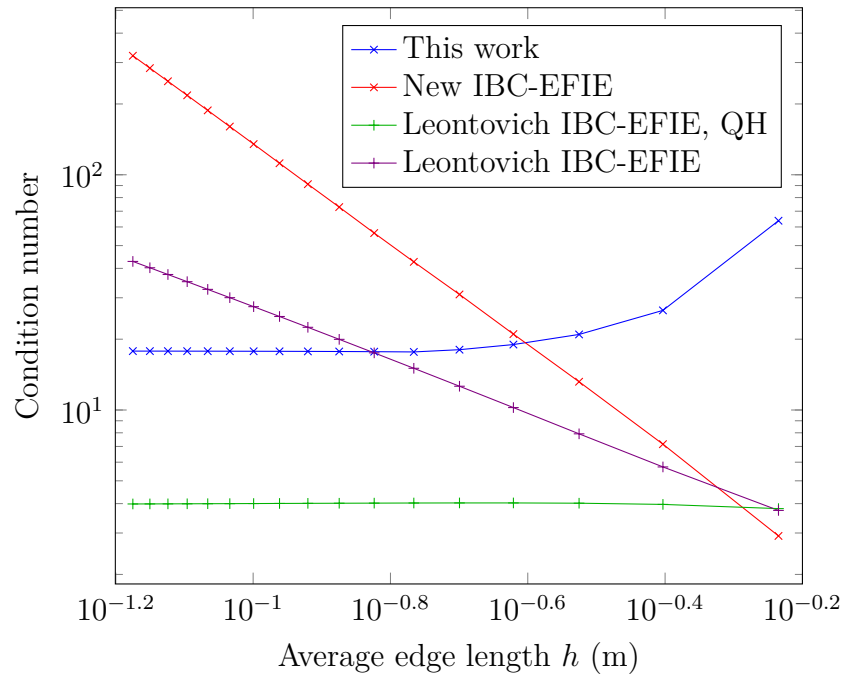


(a) Condition number as a function of the frequency

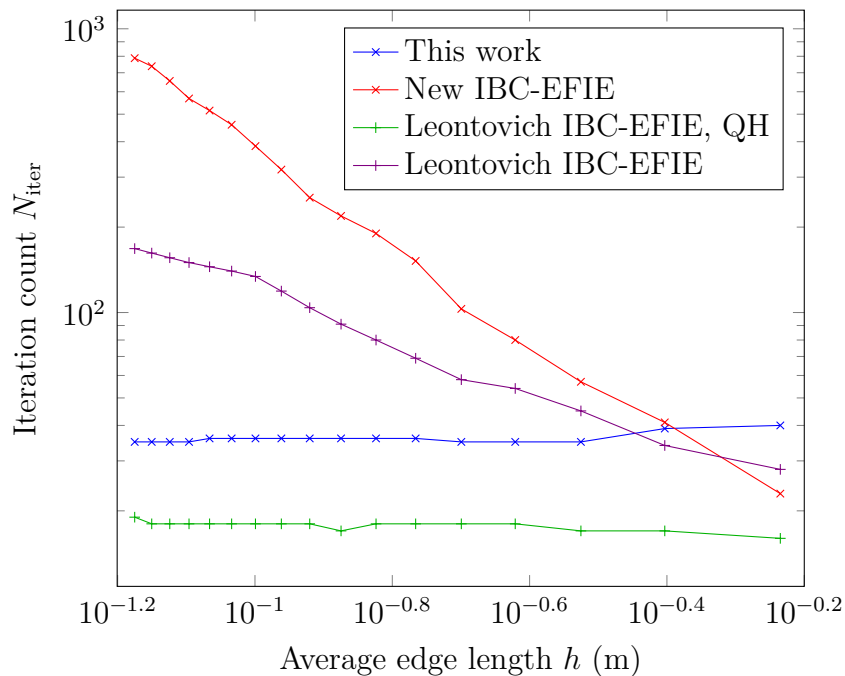


(b) Iteration count as a function of the frequency

Figure 3.4: Condition number and iteration count as functions of the frequency f when the impedance z^{imp} varies with the frequency



(a) Condition number function of the edge length



(b) Iteration count as a function of the edge length

Figure 3.5: Condition number and iteration count as functions of the edge length h

3. Multiplicative preconditioner for a new IBC-EFIE

Formulation	Condition number	Iteration count
This work	3.3×10^2	802
New IBC-EFIE	4.7×10^3	> 11199
Leontovich IBC-EFIE, QH	9.8×10^2	540
Leontovich IBC-EFIE	1.8×10^3	5989

Figure 3.6: Condition number and iteration count to solve the aircraft problem

Finally, the figure 3.5 shows the condition number and the number of iterations as functions of the average edge length h (unit sphere, $N_s = 1920$, $\delta = 0.15$ m, $z^{\text{imp}} = (0.9 + 0.9i)\eta$, $f = 50$ MHz). As it can be observed, the condition number and the number of iterations are stable as $h \rightarrow 0$. Note that the condition number of the non-preconditioned IBC-EFIE grows differently for the new IBC and the Leontovich IBC. In the case of the Leontovich IBC, the solenoidal part of the spectrum is dominated by an identity and therefore the singular values tend to a non-zero constant (see (2.34a)). However, in the case of the new IBC, there is a contribution from \mathcal{T}_s that scales as h and a contribution from \mathcal{N}_δ that also scales as h , and therefore this part of the spectrum scales proportionally to h . With the non-solenoidal part of the spectrum that scales as $1/h$, it results in a condition number that scales proportionally to $1/h^2$ for the new IBC-EFIE as opposed to a $1/h$ scaling for the Leontovich IBC-EFIE.

3.6.3 Realistic application

A comparison of the Leontovich IBC and the new IBC is done on an stealth aircraft model ($N_s = 11199$, $h \approx 0.29$ m). The excitation is a vertically polarized plane wave that is coming from the front of the aircraft ($\hat{\mathbf{p}} = \hat{\mathbf{z}}$, $\hat{\mathbf{k}} = -\hat{\mathbf{x}}$ and $E_0 = 1$ V/m in (1.42)), $z^{\text{imp}} = \eta$, $f = 30$ MHz and $\delta = 0.4$ m. The condition numbers and iteration counts have been reported in the table 3.6. In the figure 3.7, the RCS has been computed on the xy plane. As expected, there are notable differences in the RCS because the IBCs are different, even though the overall trend is similar. Decreasing δ makes the RCS computed with the new IBC to converge to the RCS computed with the Leontovich IBC. In the figure 3.8 and 3.9, the surface current densities have been represented by computing the norm of the real part of the currents $|\Re(\mathbf{J}_s)|$ in the center of each face and then interpolated. Note that the aircraft is a few wavelengths long, so at $f = 30$ MHz the low frequency preconditioning is not used to its full potential (and it might even be harmful in some cases).

3.7 Conclusion and future work

In this work, a new impedance boundary condition $\mathbf{M}_s = -z^{\text{imp}}\mathcal{N}_\delta\mathbf{J}_s$ has been introduced. It is used in the context of the EFIE instead of the classical Leontovich IBC $\mathbf{M}_s = -z^{\text{imp}}\hat{\mathbf{n}} \times \mathbf{J}_s$. The new IBC is designed to leave the solution of the problem unchanged up to a level of accuracy controllable with the length δ , and at the same time improve on the computational properties of the IBC-EFIE. Indeed, it

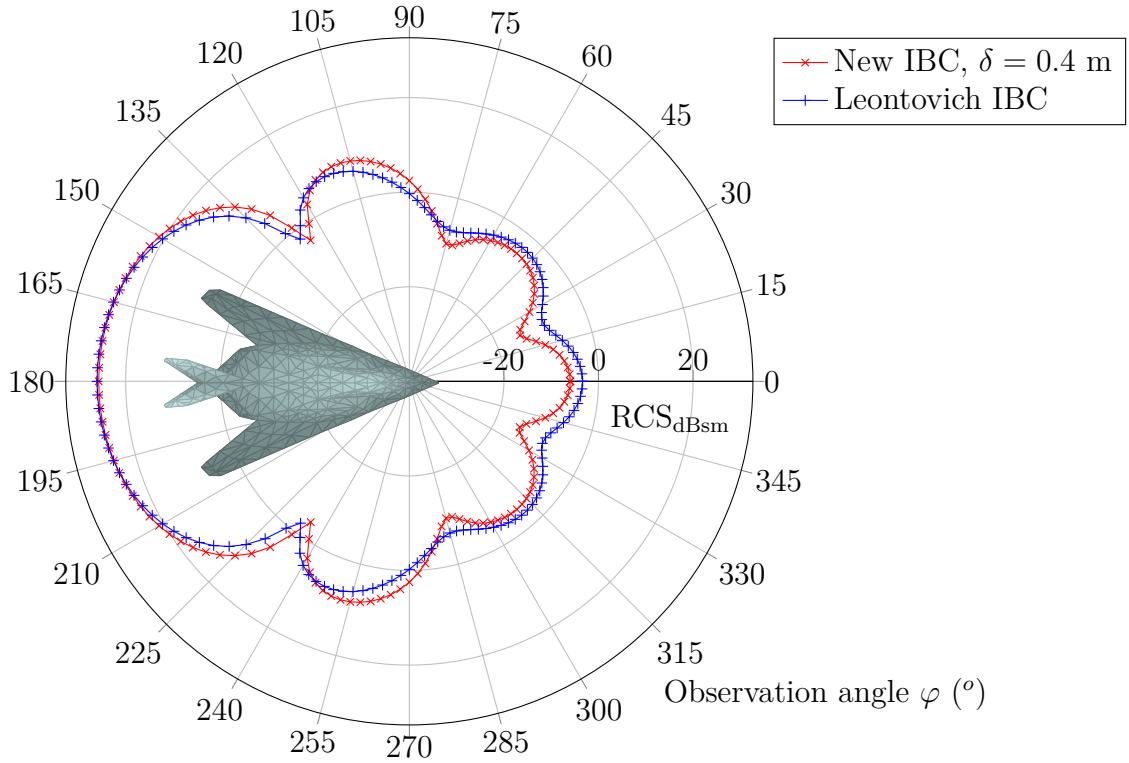
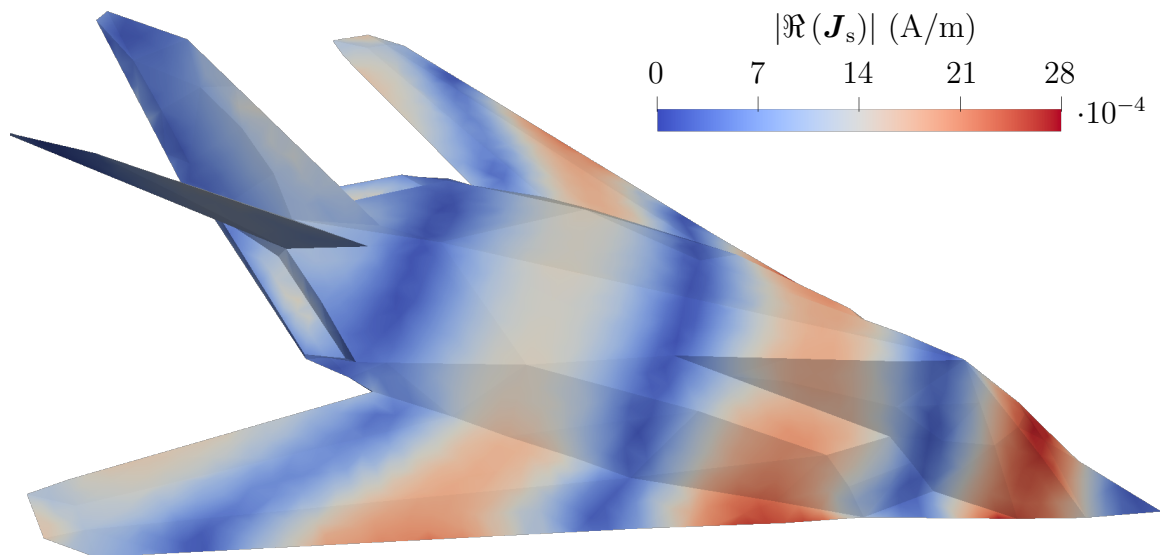


Figure 3.7: RCS (in dBsm) of an aircraft at $f = 30$ MHz with $z^{\text{imp}} = \eta$

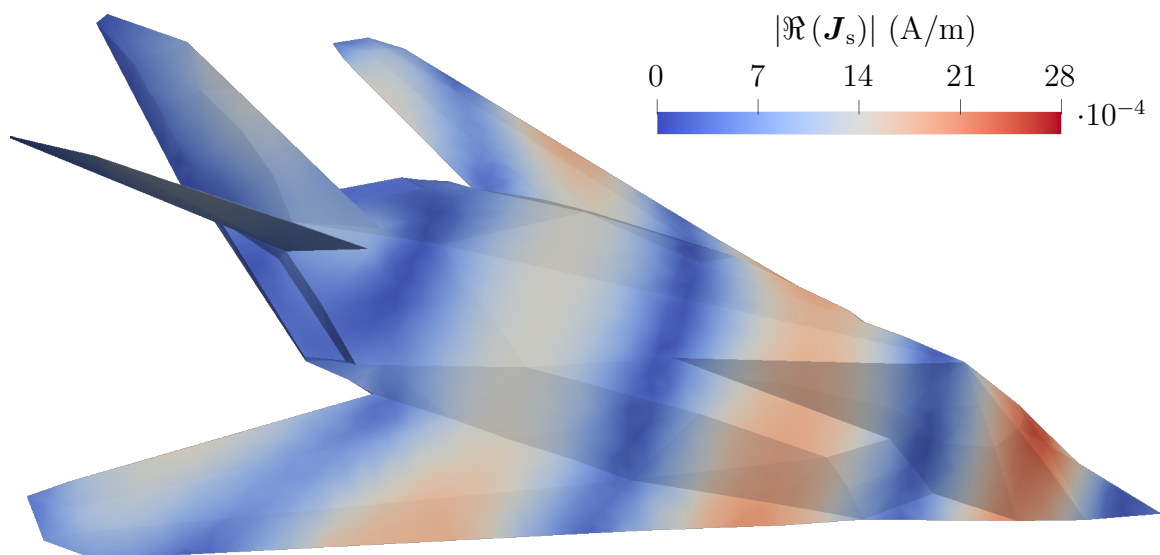
has been established through an analytic analysis on a sphere and through numerical results that this new IBC enables the IBC-EFIE to be stabilized at low frequency and for dense mesh using a multiplicative preconditioner of type Calderón. This new formulation actually reduces to the standard Calderón preconditioned PEC-EFIE in the limit $z^{\text{imp}} = 0$. This distinguishes from the Leontovich IBC-EFIE presented in the previous chapter where the dense mesh preconditioner is not compatible with the PEC-EFIE. Furthermore, the fact that this formulation does not leverage the quasi-Helmholtz decomposition to stabilize the system makes it arguably easier to comprehend and to implement.

Future work on this new IBC include the following topics:

- Spurious resonances in frequency. In the high frequency regime, the resonances can be solved by combining the IBC-EFIE with the IBC-MFIE. Preliminary results suggest that the IBC-MFIE is automatically preconditioned for the dense mesh regime using this new IBC because the unbounded part of the spectrum dominated by \mathcal{T}_h (that scales proportionally to l) is regularized by the operator \mathcal{N}_δ (that scales proportionally to $1/l$).
- Error analysis. A quantification of the difference between the classical Leontovich IBC and the new IBC depending on δ , as well as a quantification of the effect of δ on the preconditioning can be done to have a better understanding on the fine tuning of the parameter δ .

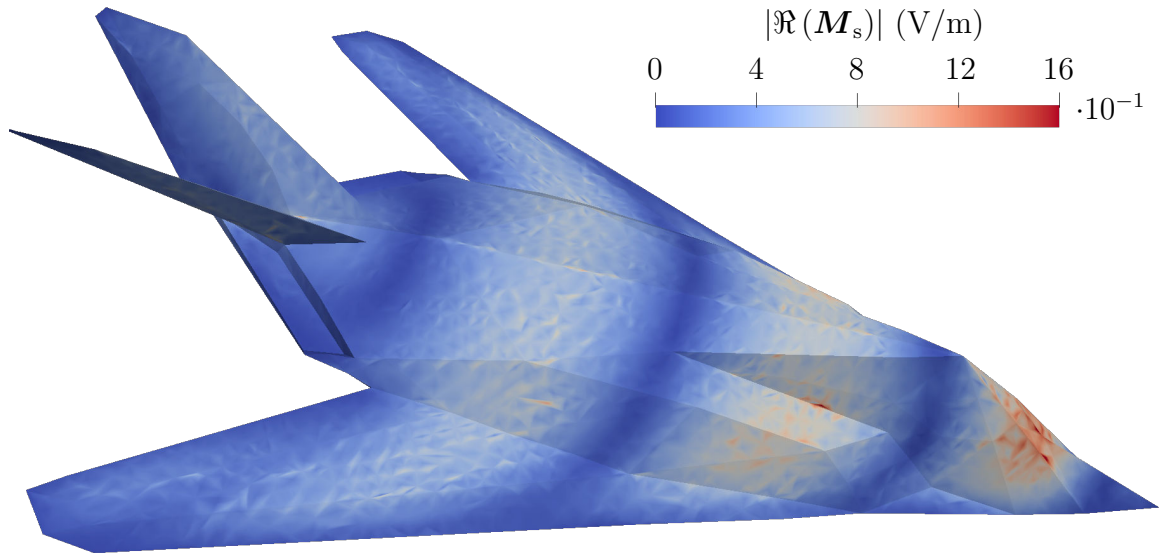


(a) Surface electric current density with new IBC ($\delta = 0.4$ m)

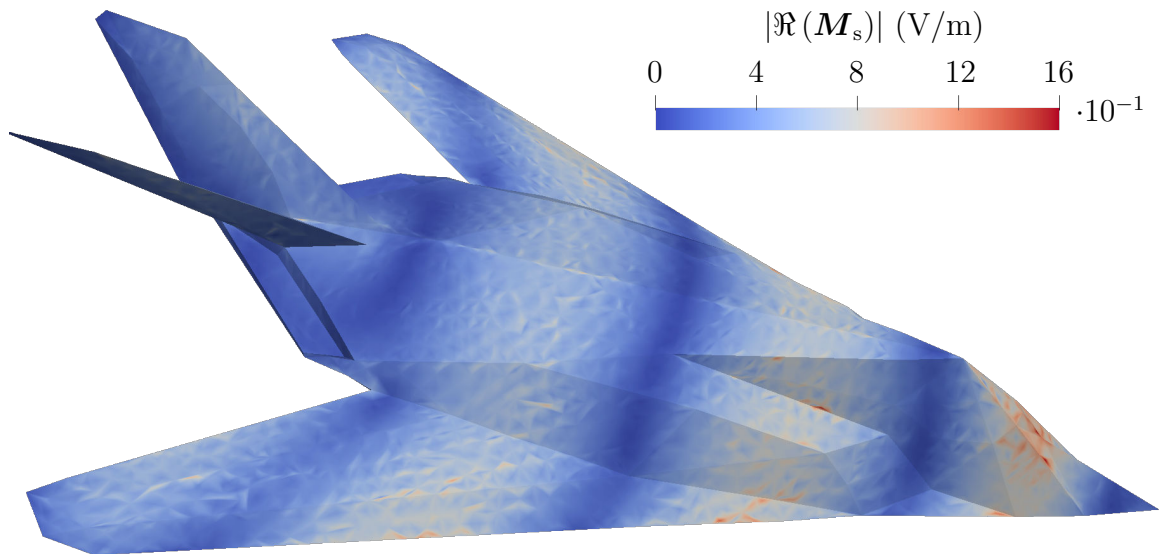


(b) Surface electric current density with the Leontovich IBC

Figure 3.8: Norm of the real part of the surface electric current density $|\Re(\mathbf{J}_s)|$ (A/m) on an aircraft



(a) Surface magnetic current density with new IBC ($\delta = 0.4$ m)



(b) Surface magnetic current density with the Leontovich IBC

Figure 3.9: Norm of the real part of the surface magnetic current density $|\Re(\mathbf{M}_s)|$ (V/m) on an aircraft

Part II

Time domain

Chapter 4

Introduction to the time domain

As opposed to the previous sections that cover the frequency domain, this section is dedicated exclusively to the time domain. Some results from the frequency domain are reused, either to avoid redoing the derivations, or to compare as analogy. The Marching-On-in-Time (MOT) scheme is introduced through the classical testing procedure with temporal basis functions to discretize the equations in time but without extensive details. Rather in this work, the time discretization and the MOT scheme are obtained with a convolution quadrature using implicit Runge-Kutta methods. Finally, the DC instability that has no counterpart in the frequency domain is also presented.

4.1 Surface integral equations in the time domain

The integral equations can be derived from the Maxwell's equations in the time domain (1.1), with derivations that are analogue to the frequency domain. However, to take advantage of the previous chapters, an inverse Fourier transform is used in the following to retrieve the main results. Note that in general, the inverse Fourier transform of frequency domain quantities is not sufficient for all intents and purposes in the time domain. The inverse Fourier transform \mathcal{F}^{-1} is defined by

$$\mathcal{F}^{-1}(X)(t) = \frac{1}{2\pi} \int_{\omega \in \mathbb{R}} X(\omega) e^{i\omega t} d\omega. \quad (4.1)$$

In particular, the following identities can be used to transform the results from the frequency domain to the time domain

$$\mathcal{F}^{-1}(i\omega X)(t) = \frac{\partial}{\partial t} (\mathcal{F}^{-1}(X)(t)) \quad (4.2a)$$

$$\mathcal{F}^{-1}\left(\frac{1}{i\omega} X\right)(t) = \int_{t'=-\infty}^t \mathcal{F}^{-1}(X)(t') dt' \quad (4.2b)$$

$$\mathcal{F}^{-1}(e^{-i\omega\tau} X)(t) = \mathcal{F}^{-1}(X)(t - \tau), \quad (4.2c)$$

4. Introduction to the time domain

or equivalently using the wave number $k = \frac{\omega}{c}$ where c is the speed of light in the medium

$$\mathcal{F}^{-1}(ikX)(t) = \frac{1}{c} \frac{\partial}{\partial t} (\mathcal{F}^{-1}(X)(t)) \quad (4.3a)$$

$$\mathcal{F}^{-1}\left(\frac{1}{ik}X\right)(t) = c \int_{t'=-\infty}^t \mathcal{F}^{-1}(X)(t') dt' \quad (4.3b)$$

$$\mathcal{F}^{-1}(e^{-ikR}X)(t) = \mathcal{F}^{-1}(X)\left(t - \frac{R}{c}\right). \quad (4.3c)$$

Given the source currents $\mathbf{J}(\mathbf{r}, t)$ and $\mathbf{M}(\mathbf{r}, t)$, and the source charges $\rho(\mathbf{r}, t)$ and $\rho_m(\mathbf{r}, t)$, the generalized (in the sense that they include the non-physical magnetic charge ρ_m and current \mathbf{M} densities) Maxwell's equations that are the counterparts of (1.29) read

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \rho(\mathbf{r}, t) \quad (4.4a)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = \rho_m(\mathbf{r}, t) \quad (4.4b)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} - \mathbf{M}(\mathbf{r}, t) \quad (4.4c)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} + \mathbf{J}(\mathbf{r}, t). \quad (4.4d)$$

In the time domain, the retarded potentials that correspond to the frequency domain potentials (1.27) and (1.31) are

$$\mathbf{A}(\mathbf{r}, t) = \mu \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \mathbf{J}\left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right) dV' \quad (4.5a)$$

$$\phi(\mathbf{r}, t) = \frac{1}{\varepsilon} \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \rho\left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right) dV' \quad (4.5b)$$

$$\mathbf{F}(\mathbf{r}, t) = \varepsilon \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \mathbf{M}\left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right) dV' \quad (4.5c)$$

$$\phi_m(\mathbf{r}, t) = \frac{1}{\mu} \iiint_{\mathbf{r}' \in \mathbb{R}^3} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \rho_m\left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right) dV' \quad (4.5d)$$

with the Lorenz gauge (see (1.22) and (1.32))

$$\nabla \cdot \mathbf{A}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial}{\partial t} \phi(\mathbf{r}, t) \quad (4.6)$$

$$\nabla \cdot \mathbf{F}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial}{\partial t} \phi_m(\mathbf{r}, t). \quad (4.7)$$

The potentials verify the wave equation which is the time dependent counterpart of

the Helmholtz equation

$$\nabla^2 \mathbf{A}(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{A}(\mathbf{r}, t) = -\mu \mathbf{J}(\mathbf{r}, t) \quad (4.8a)$$

$$\nabla^2 \phi(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \phi(\mathbf{r}, t) = -\frac{\rho(\mathbf{r}, t)}{\varepsilon} \quad (4.8b)$$

$$\nabla^2 \mathbf{F}(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{F}(\mathbf{r}, t) = -\varepsilon \mathbf{M}(\mathbf{r}, t) \quad (4.8c)$$

$$\nabla^2 \phi_m(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \phi_m(\mathbf{r}, t) = -\frac{\rho_m(\mathbf{r}, t)}{\mu}. \quad (4.8d)$$

The scattered fields are (see (1.34))

$$\mathbf{E}^{\text{sca}}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial}{\partial t} \mathbf{A}(\mathbf{r}, t) - \nabla \phi(\mathbf{r}, t) - \frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}, t) \quad (4.9a)$$

$$\mathbf{H}^{\text{sca}}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial}{\partial t} \mathbf{F}(\mathbf{r}, t) - \nabla \phi_m(\mathbf{r}, t) + \frac{1}{\mu} \nabla \times \mathbf{A}(\mathbf{r}, t), \quad (4.9b)$$

or in terms of currents (see (1.35))

$$\mathbf{E}^{\text{sca}}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial}{\partial t} \mathbf{A}(\mathbf{r}, t) + c \int_{t'=-\infty}^t \nabla \nabla \cdot \mathbf{A}(\mathbf{r}, t') dt' - \frac{1}{\varepsilon} \nabla \times \mathbf{F}(\mathbf{r}, t) \quad (4.10a)$$

$$\mathbf{H}^{\text{sca}}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial}{\partial t} \mathbf{F}(\mathbf{r}, t) + c \int_{t'=-\infty}^t \nabla \nabla \cdot \mathbf{F}(\mathbf{r}, t') dt' + \frac{1}{\mu} \nabla \times \mathbf{A}(\mathbf{r}, t). \quad (4.10b)$$

Now, considering a scatterer whose boundary is Γ , the boundary conditions at a discontinuous interface derived in section 1.2.6 are also valid in the time domain. So, there are surface equivalent currents $\mathbf{J}_s(\mathbf{r}, t)$ and $\mathbf{M}_s(\mathbf{r}, t)$ that verify the boundary conditions (1.57). In particular, in the exterior domain the vector potentials are

$$\mathbf{A}(\mathbf{r}, t) = \mu \iint_{\mathbf{r}' \in \Gamma} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \mathbf{J}_s \left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c} \right) dS' \quad (4.11a)$$

$$\mathbf{F}(\mathbf{r}, t) = \varepsilon \iint_{\mathbf{r}' \in \Gamma} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \mathbf{M}_s \left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c} \right) dS'. \quad (4.11b)$$

Using the surface equivalence principle to set the fields in the interior domain to $\mathbf{0}$ as is done in section 1.2.7.1, the boundary conditions (1.58) become for all $\mathbf{r} \in \Gamma$ and $t \in \mathbb{R}$,

$$\mathbf{J}_s(\mathbf{r}, t) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}(\mathbf{r}, t) \quad (4.12a)$$

$$\mathbf{M}_s(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) \times \hat{\mathbf{n}}(\mathbf{r}). \quad (4.12b)$$

Then, writing the total fields as sums of the incident fields and the scattered fields

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}^{\text{inc}}(\mathbf{r}, t) + \mathbf{E}^{\text{sca}}(\mathbf{r}, t) \quad (4.13a)$$

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}^{\text{inc}}(\mathbf{r}, t) + \mathbf{H}^{\text{sca}}(\mathbf{r}, t) \quad (4.13b)$$

and taking the limit when \mathbf{r} approaches Γ results in the time domain Electric Field Integral Equation (EFIE) and the time domain Magnetic Field Integral Equation (MFIE)

$$\eta \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{M}_s(\mathbf{r}, t) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, t) \quad (4.14a)$$

$$-\frac{1}{\eta} \mathcal{T} \mathbf{M}_s(\mathbf{r}, t) + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{J}_s(\mathbf{r}, t) = \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{H}^{\text{inc}}(\mathbf{r}, t) \quad (4.14b)$$

where the time domain surface integral operators are defined as (see their frequency domain counterparts (1.68a), (1.69) and (1.70))

$$(\mathcal{T} \mathbf{f})(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial (\mathcal{T}_s \mathbf{f})(\mathbf{r}, t)}{\partial t} + c \int_{t'=-\infty}^t (\mathcal{T}_h \mathbf{f})(\mathbf{r}, t') dt' \quad (4.15a)$$

$$(\mathcal{T}_s \mathbf{f})(\mathbf{r}, t) = \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \mathbf{f}\left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right) dS' \quad (4.15b)$$

$$(\mathcal{T}_h \mathbf{f})(\mathbf{r}, t) = \hat{\mathbf{n}}(\mathbf{r}) \times \nabla \iint_{\mathbf{r}' \in \Gamma} \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \nabla' \cdot \mathbf{f}\left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right) dS' \quad (4.15c)$$

$$(\mathcal{K} \mathbf{f})(\mathbf{r}, t) = \hat{\mathbf{n}}(\mathbf{r}) \times \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla \left(\frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \right) \times \mathbf{f}\left(\mathbf{r}', t - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right) dS' \quad (4.15d)$$

$$(\mathcal{I} \mathbf{f})(\mathbf{r}, t) = \mathbf{f}(\mathbf{r}, t). \quad (4.15e)$$

It is worth noting that the EFIE and MFIE are usually time differentiated, mainly to get rid of the time integration in the definition of \mathcal{T} in (4.15a).

4.2 Time domain discretization and marching-on-in-time

The differentiated PEC-EFIE is considered in the following

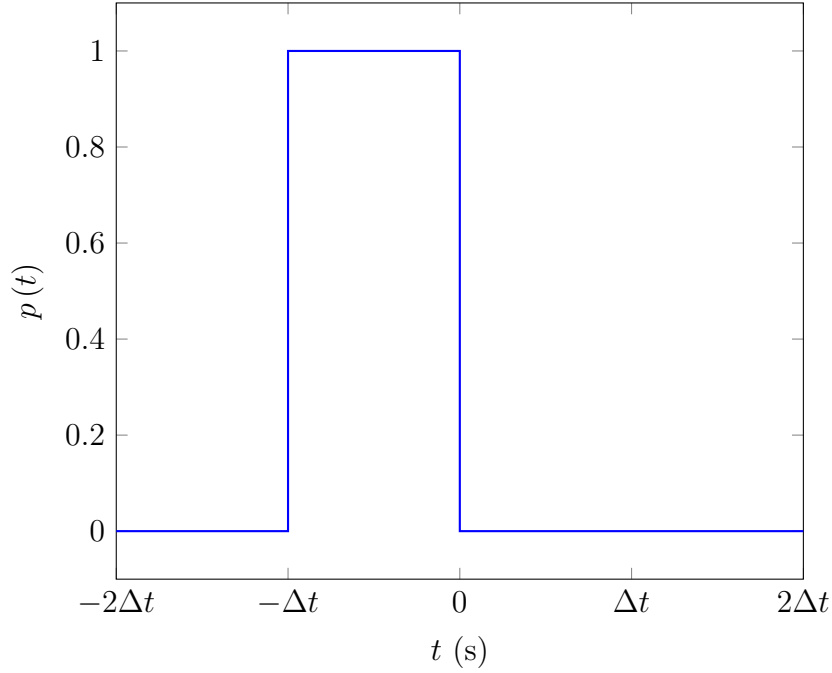
$$\eta \frac{\partial}{\partial t} \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) = -\hat{\mathbf{n}}(\mathbf{r}) \times \frac{\partial}{\partial t} \mathbf{E}^{\text{inc}}(\mathbf{r}, t). \quad (4.16)$$

As in the frequency domain, the RWG basis functions \mathbf{f}_n (1.90) are a common choice for the spacial discretization. So, the current is approximated as

$$\mathbf{J}_s(\mathbf{r}, t) \approx \sum_{n=1}^{N_s} [\mathbf{J}(t)]_n \mathbf{f}_n(\mathbf{r}). \quad (4.17)$$

Compared to the frequency domain, a notable difference is the time dependency of the coefficients vector $\mathbf{J}(t)$. With a Galerkin scheme (i.e. using the same functions for the sources and testings), the space discretized system has the form

$$(\mathbf{Z} * \mathbf{J})(t) = \mathbf{E}(t) \quad (4.18)$$


 Figure 4.1: Pulse basis function p

where for all $m \in [1, N_s]$,

$$[(\mathbf{Z} * \mathbf{J})(t)]_m = \eta \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) \, dS \quad (4.19a)$$

$$[\mathbf{E}(t)]_m = - \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathbf{E}^{\text{inc}}(\mathbf{r}, t) \, dS. \quad (4.19b)$$

The expression of $\mathbf{Z}(t)$ is not needed in this section, but it can be found in (4.72a) in section 4.3.4.

It is usually assumed that there exists a time t_0 such that all the fields in the neighbourhood of Γ and the currents vanish for all $t < t_0$. Similarly, there is a time t_{N_t} for which the fields vanish for all $t > t_{N_t}$ (transient scattering), or if this time is too long, t_{N_t} is chosen such that the solution is not of interest after that time. Then, the solution is computed for $t \in [t_0, t_{N_t}]$. This range is subdivided into N_t time steps that have a duration $\Delta t > 0$. In other words, the solution is computed for all $t \in [t_0, t_0 + N_t \Delta t]$. The time stamps are defined as

$$t_i = t_0 + i \Delta t. \quad (4.20)$$

Then, the temporal basis functions are introduced. A common choice for the sources are the pulses $(p_i)_{i=1}^{N_t}$. The same pulse function p is translated to build each function

$$p_i(t) = p(t - t_i) \quad (4.21)$$

where p is illustrated in the figure 4.1 and defined as

$$p(t) = \begin{cases} 1 & \text{if } t \in [-\Delta t, 0] \\ 0 & \text{otherwise.} \end{cases} \quad (4.22)$$

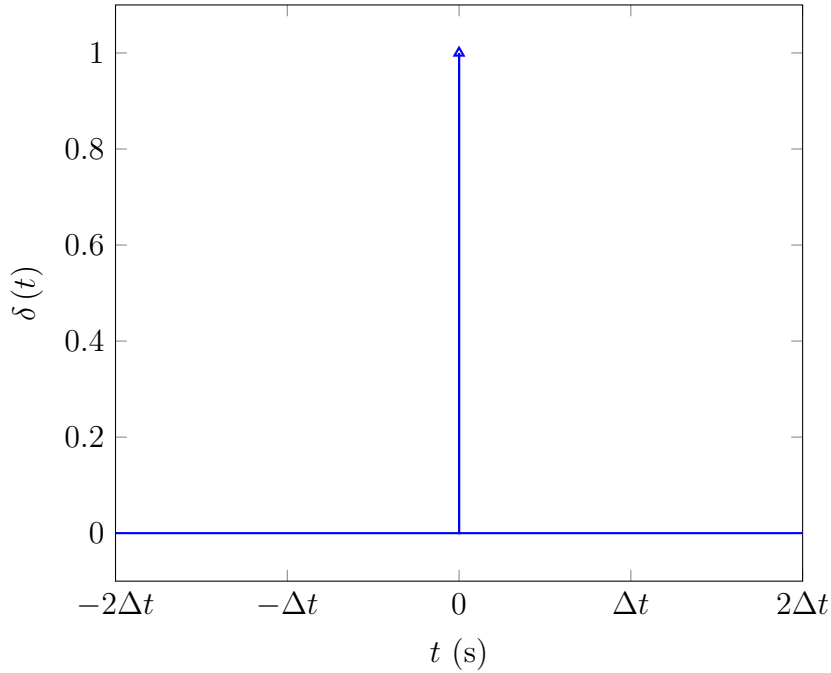


Figure 4.2: Dirac delta distribution δ

The current coefficients $[\mathbf{J}(t)]_n$ ($n \in [1, N_s]$) can be approximated by

$$[\mathbf{J}(t)]_n \approx \sum_{j=1}^{N_t} [\mathbf{J}_j]_n p(t - t_j) \quad (4.23)$$

and the current is

$$\mathbf{J}_s(\mathbf{r}, t) \approx \sum_{n=1}^{N_s} \sum_{j=1}^{N_t} [\mathbf{J}_j]_n \mathbf{f}_n(\mathbf{r}) p(t - t_j). \quad (4.24)$$

The discretization of the current can be interpreted as a sequence of N_t arrays of N_s coefficients $(\mathbf{J}_j)_{j=1}^{N_t} \in (\mathbb{R}^{N_s})^{N_t}$.

The inner product between two integrable real functions f and g is defined as

$$\langle f, g \rangle = \int_{t \in \mathbb{R}} f(t)g(t)dt. \quad (4.25)$$

In the following x denotes a function that is used to build the temporal testing functions $(x_i)_{i=1}^{N_t}$ where $x_i(t) = x(t - t_i)$. A common choice is the point testing, which corresponds to choosing $x = \delta$. The Dirac delta distribution δ (illustrated in the figure 4.2) verifies $\langle \delta, f \rangle = f(0)$ for any integrable function f . Another common choice for the temporal testing is to use a Galerkin scheme. In this case it corresponds to choosing the pulses ($x = p$) for the testing basis functions. When it is combined with a Galerkin scheme in space it is called space-time Galerkin. Testing the RHS

of (4.18) in time for all $i \in [1, N_t]$ results in a sequence of vectors $(\mathbf{E}_i)_{i=1}^{N_t}$ whose coefficients are $(m \in [1, N_s])$

$$[\mathbf{E}_i]_m = \langle x_i, [\mathbf{E}(t)]_m \rangle \quad (4.26a)$$

$$= - \int_{t \in \mathbb{R}} x_i(t) \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathbf{E}^{\text{inc}}(\mathbf{r}, t) \, dSdt \quad (4.26b)$$

$$= - \int_{t \in \mathbb{R}} x(t) \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathbf{E}^{\text{inc}}(\mathbf{r}, t + i\Delta t) \, dSdt. \quad (4.26c)$$

Testing the LHS of (4.18) in time results in

$$\langle x_i, [(\mathbf{Z} * \mathbf{J})(t)]_m \rangle = \eta \int_{t \in \mathbb{R}} x_i(t) \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) \, dSdt \quad (4.27a)$$

$$= \eta \sum_{n=1}^{N_s} \sum_{j=1}^{N_t} [\mathbf{J}_j]_n \int_{t \in \mathbb{R}} x_i(t) \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathcal{T}(\mathbf{f}_{np_j})(\mathbf{r}, t) \, dSdt \quad (4.27b)$$

$$= \eta \sum_{n=1}^{N_s} \sum_{j=1}^{N_t} [\mathbf{J}_j]_n \int_{t \in \mathbb{R}} x(t) \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathcal{T}(\mathbf{f}_{np})(\mathbf{r}, t + (i-j)\Delta t) \, dSdt. \quad (4.27c)$$

By introducing the sequence of matrices $(\mathbf{Z}_j)_j$ defined by

$$[\mathbf{Z}_j]_{mn} = \eta \int_{t \in \mathbb{R}} x(t) \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathcal{T}(\mathbf{f}_{np})(\mathbf{r}, t + j\Delta t) \, dSdt, \quad (4.28)$$

the fully discretized system is (for all $i \in [1, N_t]$)

$$\sum_{j=1}^{N_t} \mathbf{Z}_{i-j} \mathbf{J}_j = \mathbf{E}_i. \quad (4.29)$$

In general, the temporal testing function x must be chosen to ensure causality such that $\mathbf{Z}_j = \mathbf{0}$ for all $j < 0$. With this assumption, (4.29) becomes

$$\sum_{j=1}^i \mathbf{Z}_{i-j} \mathbf{J}_j = \mathbf{E}_i \quad (4.30)$$

or equivalently,

$$\sum_{j=0}^{i-1} \mathbf{Z}_j \mathbf{J}_{i-j} = \mathbf{E}_i. \quad (4.31)$$

This expression (4.31) can be rewritten for each $i \in [1, N_t]$ as

$$\mathbf{Z}_0 \mathbf{J}_i = \mathbf{E}_i - \sum_{j=1}^{i-1} \mathbf{Z}_j \mathbf{J}_{i-j}. \quad (4.32)$$

This equation (4.32) corresponds to the so-called Marching-On-in-Time (MOT) that solves the unknown vectors \mathbf{J}_i sequentially for each time step i . This is done by first computing the RHS of (4.32) using the vectors \mathbf{J}_j known for each $j < i$ and then by solving the linear system involving \mathbf{Z}_0 . In the case of the EFIE, it can be seen from (4.15) and (4.28) that $\mathbf{Z}_j = \mathbf{0}$ when $j > \max_{\mathbf{r}, \mathbf{r}' \in \Gamma} \frac{|\mathbf{r} - \mathbf{r}'|}{c\Delta t}$ (up to an additional constant term that depends on the support of the temporal basis functions). So there is a constant N_{conv} such that

$$\mathbf{Z}_0 \mathbf{J}_i = \mathbf{E}_i - \sum_{j=1}^{N_{\text{conv}}} \mathbf{Z}_j \mathbf{J}_{i-j}. \quad (4.33)$$

In the above expression, it is assumed that $\mathbf{J}_{i-j} = \mathbf{0}$ if $i \leq j$ in the sum.

The computational complexity of computing (4.32) is $\mathcal{O}(N_t N_s^2)$. Fortunately there are algorithms to speed up the computation such as the Time Domain Adaptive Integral Method (TD-AIM) [43] whose computational complexity is $\mathcal{O}(N_t N_s^{1.5} \log(N_s)^2)$, the Plane Wave Time Domain (PWTD) algorithm [44] that computes the RHS with $\mathcal{O}(N_t N_s^{1.5} \log(N_s))$ operations, and the multi-level PWTD algorithm [45] that reduces it to $\mathcal{O}(N_t N_s \log(N_s)^2)$.

4.3 Implicit Runge-Kutta convolution quadrature

Time testing with temporal basis functions is not the only way to discretize a time domain integral equation. Another technique is the convolution quadrature that basically consists in first transforming the equation to Laplace domain. Then, the Laplace domain equation is transformed into an equation in Z -domain using Implicit Runge-Kutta methods. And finally, an inverse Z -transform yields a fully discretized system in time. As in the previous section, this procedure results in a MOT scheme. Again, the differentiated PEC-EFIE is used to introduce the technique.

4.3.1 Laplace transform and Z -transform

4.3.1.1 Laplace transform

The Laplace transform transforms a function X in time domain to a function $\mathcal{L}(X)$ in Laplace domain. It is defined for $s \in \mathbb{C}$ in the region of convergence as

$$\mathcal{L}(X)(s) = \int_{t=0}^{+\infty} X(t) e^{-st} dt. \quad (4.34)$$

To simplify the notations, the Laplace transform of X is also noted X , such that $X(s) = \mathcal{L}(X)(s)$ and the distinction from the time domain function can be done by looking at the argument. The inverse Laplace transform is

$$X(t) = \mathcal{L}^{-1}(X)(t) = \frac{1}{2\pi i} \int_{s=\sigma-i\infty}^{\sigma+i\infty} X(s) e^{st} ds \quad (4.35)$$

where the integration is done in the region of convergence on the line $\Re(s) = \sigma$ with all the singularities of $X(s)$ on the left of it.

4.3.1.2 Z-transform

The Z-transform transforms a sequence $(X_n)_{n \in \mathbb{N}}$ in discrete time domain to a function $\mathcal{Z}(X)$ in Z-domain defined for $z \in \mathbb{C}$ in the region of convergence as

$$\mathcal{Z}(X)(z) = \sum_{n=0}^{+\infty} X_n z^{-n}. \quad (4.36)$$

Again, if there is no possible confusion, $\mathcal{Z}(X)(z)$ is simply noted $X(z)$. The inverse Z-transform of X is a sequence $(\mathcal{Z}^{-1}(X)_n)_{n \in \mathbb{N}}$ defined as

$$X_n = \mathcal{Z}^{-1}(X)_n = \frac{1}{2\pi i} \oint_{z \in \mathcal{C}} X(z) z^{n-1} dz \quad (4.37)$$

where \mathcal{C} is a counterclockwise contour in the region of convergence of X and around the origin. By choosing \mathcal{C} to be a circle of radius ρ , an approximation of the inverse Z-transform is obtained with the used a trapezoidal rule on Q subintervals that subdivide $[-\pi, \pi]$,

$$X_n = \mathcal{Z}^{-1}(X)_n = \frac{\rho^n}{2\pi} \int_{\theta=-\pi}^{\pi} X(\rho e^{i\theta}) e^{i\theta n} d\theta \quad (4.38a)$$

$$\approx \frac{\rho^n}{Q} \sum_{q=0}^{Q-1} X(\rho e^{2\pi i \frac{q}{Q}}) e^{2\pi i \frac{q}{Q} n}. \quad (4.38b)$$

Notable properties of the Z-transform are

$$\mathcal{Z}((X_{n-m})_n)(z) = z^{-m} \mathcal{Z}(X)(z) \quad (4.39a)$$

$$\mathcal{Z}((X_{n+1})_n)(z) = z \mathcal{Z}(X)(z) - z X_0 \quad (4.39b)$$

$$\mathcal{Z}((\delta_{n,0})_n)(z) = 1 \quad (4.39c)$$

$$\mathcal{Z}(X * Y) = \mathcal{Z} \left(\left(\sum_{m=0}^n X_m Y_{n-m} \right)_n \right) = \mathcal{Z}(X) \mathcal{Z}(Y). \quad (4.39d)$$

The Z-transform can be thought as the discrete counterpart of the Laplace transform by considering the following relation where x is a function of time, X is the function x sampled with Dirac deltas and Y is a sequence of samples of x

$$X(t) = x(t) \sum_{n=0}^{+\infty} \delta(t - n\Delta t) \quad (4.40a)$$

$$Y_n = x(n\Delta t) \quad (4.40b)$$

$$\mathcal{L}(X)(s) = \mathcal{Z}(Y)(e^{s\Delta t}). \quad (4.40c)$$

4.3.2 Runge-Kutta methods

The Runge-Kutta methods are a class of techniques to solve differential equations in the form

$$\frac{dy(t)}{dt} = F(t, y) \quad (4.41a)$$

$$y(t_0) = y_0. \quad (4.41b)$$

where the unknown y is a function of time t whose initial value y_0 at t_0 is known and whose derivative is a function of t and y . The solution is computed consecutively at each time $t_i = t_0 + i\Delta t$ and is noted $y_i = y(t_i)$. At the time step i , the Runge-Kutta methods compute the solution for the next time step y_{i+1} by adding to the current solution y_i a weighted sum of p interpolants $[\mathbf{F}_i]_k$ of the slope

$$y_{i+1} = y_i + \Delta t \sum_{k=1}^p \mathbf{b}_k [\mathbf{F}_i]_k, \quad (4.42)$$

where p is the number of stages, $[\mathbf{F}_i]_k$ ($k \in [1, p]$) are the interpolants of the slope and $\mathbf{b} \in \mathbb{R}^p$ is an array of weights. The $[\mathbf{F}_i]_k$ are evaluated at an intermediate time $t_i + \mathbf{c}_k \Delta t$ that is determined by the coefficients in the array $\mathbf{c} \in \mathbb{R}^p$. They are evaluated as

$$[\mathbf{F}_i]_k = F(t_i + \mathbf{c}_k \Delta t, [\mathbf{Y}_i]_k) \quad (4.43a)$$

$$[\mathbf{Y}_i]_k = y_i + \Delta t \sum_{l=1}^p \mathbf{A}_{kl} [\mathbf{F}_i]_l \quad (4.43b)$$

where $\mathbf{A} \in \mathbb{R}^{p \times p}$. If $\mathbf{A}_{kl} = 0$ for all $l \geq k$ i.e. if each $[\mathbf{F}_i]_k$ only depends on the previous $[\mathbf{F}_i]_l$, then the method is said to be explicit and the interpolants $[\mathbf{F}_i]_k$ can be computed sequentially. Otherwise, the method is implicit. In this work, only Implicit Runge-Kutta (IRK) methods are considered.

The Runge-Kutta methods are therefore specified by the matrix $\mathbf{A} \in \mathbb{R}^{p \times p}$ and the vectors $\mathbf{b}, \mathbf{c} \in \mathbb{R}^p$ that are commonly displayed in a Butcher tableau [46]

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^\top \end{array} \quad (4.44)$$

An example of such an IRK method is the 3 stages Radau IIA whose Butcher tableau is

$$\begin{array}{c|ccc} \frac{4 - \sqrt{6}}{10} & \frac{88 - 7\sqrt{6}}{360} & \frac{296 - 169\sqrt{6}}{1800} & \frac{-2 + 3\sqrt{6}}{225} \\ \frac{4 + \sqrt{6}}{10} & \frac{296 + 169\sqrt{6}}{1800} & \frac{88 + 7\sqrt{6}}{360} & \frac{-2 - 3\sqrt{6}}{225} \\ 1 & \frac{16 - \sqrt{6}}{36} & \frac{16 + \sqrt{6}}{36} & \frac{1}{9} \\ \hline & \frac{16 - \sqrt{6}}{36} & \frac{16 + \sqrt{6}}{36} & \frac{1}{9} \end{array} \quad (4.45)$$

Rewriting (4.42) and (4.43b) in vector form results in

$$y_{i+1} = y_i + \Delta t \mathbf{b}^\top \mathbf{F}_i \quad (4.46a)$$

$$\mathbf{Y}_i = y_i \mathbf{1} + \Delta t \mathbf{A} \mathbf{F}_i \quad (4.46b)$$

where $\mathbf{1} \in \mathbb{R}^p$ is the all-one vector. From (4.46b),

$$\mathbf{F}_i = \frac{1}{\Delta t} \mathbf{A}^{-1} (\mathbf{Y}_i - y_i \mathbf{1}). \quad (4.47)$$

Inserting it in (4.46a) results in

$$y_{i+1} = \mathbf{b}^\top \mathbf{A}^{-1} \mathbf{Y}_i + (1 - \mathbf{b}^\top \mathbf{A}^{-1} \mathbf{1}) y_i. \quad (4.48)$$

Some Runge-Kutta methods are called L-stable [47] or strongly A-stable [48]. They have the property $\mathbf{b}^\top \mathbf{A}^{-1} = (0, \dots, 0, 1)$, in particular

$$\mathbf{b}^\top \mathbf{A}^{-1} \mathbf{1} = 1 \quad (4.49)$$

and thus, (4.48) simplifies to

$$y_{i+1} = \mathbf{b}^\top \mathbf{A}^{-1} \mathbf{Y}_i = [\mathbf{Y}_i]_p. \quad (4.50)$$

A particular example that is used in the next section is the differential equation

$$\frac{dy(t)}{dt} = sy(t) + g(t) \quad (4.51a)$$

$$y(0) = 0 \quad (4.51b)$$

where g is a known function and s is a complex parameter. Note $[\mathbf{g}_i]_k = g(t_i + \mathbf{c}_k \Delta t)$. With these notations, $F(t, y) = sy(t) + g(t)$ and $\mathbf{F}_i = s\mathbf{Y}_i + \mathbf{g}_i$. Z-transforming (4.46) results in

$$(z - 1) \mathcal{Z}(y)(z) = \Delta t \mathbf{b}^\top \mathcal{Z}(\mathbf{F})(z) + zy_0 = \Delta t \mathbf{b}^\top \mathcal{Z}(\mathbf{F})(z) \quad (4.52a)$$

$$\mathcal{Z}(\mathbf{Y})(z) = \mathcal{Z}(y)(z) \mathbf{1} + \Delta t \mathbf{A} \mathcal{Z}(\mathbf{F})(z) \quad (4.52b)$$

where $y(0) = y_0 = 0$ was used to cancel the term zy_0 in (4.52a). In the following z is different than 1. Eventually, it implies that the contour \mathcal{C} in the inverse Z-transform (4.37) is circle with a radius $\rho \neq 1$. Thus,

$$\mathcal{Z}(\mathbf{Y})(z) = \Delta t \left(\mathbf{A} + \frac{\mathbf{1b}^\top}{z - 1} \right) \mathcal{Z}(\mathbf{k})(z) \quad (4.53a)$$

$$= \Delta t \left(\mathbf{A} + \frac{\mathbf{1b}^\top}{z - 1} \right) (s \mathcal{Z}(\mathbf{Y})(z) + \mathcal{Z}(\mathbf{g})(z)). \quad (4.53b)$$

So

$$\mathcal{Z}(\mathbf{Y})(z) = (\mathbf{s}(z) - \mathbf{s}\mathbf{1}_p)^{-1} \mathcal{Z}(\mathbf{g})(z) \quad (4.54)$$

4. Introduction to the time domain

where \mathbf{I}_p is the $p \times p$ identity matrix and

$$\mathbf{s}(z) = \frac{1}{\Delta t} \left(\mathbf{A} + \frac{\mathbf{1}\mathbf{b}^\top}{z-1} \right)^{-1}. \quad (4.55)$$

Note that with (4.49), using the Sherman-Morrison formula results in

$$\mathbf{s}(z) = \frac{1}{\Delta t} \left(\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{1}\mathbf{b}^\top\mathbf{A}^{-1}z^{-1} \right). \quad (4.56)$$

which makes the inverse Z-transforms easy to compute because a multiplication by the powers of z^{-1} corresponds to taking the previous element in the time domain sequence (see (4.39a)).

4.3.3 Convolution quadrature using IRK methods

The convolution quadrature aims to approximate the following convolution

$$x(t) = (f * g)(t) = \int_{u=0}^t f(t-u)g(u)du \quad (4.57)$$

by

$$x(j\Delta t) \approx x_j = \sum_{i=0}^j w_{j-i}g_i = (w * g)_j \quad (4.58)$$

where f and g are functions of time, $g_i = g(i\Delta t)$, and w_{j-i} are weights that are functions of the time step Δt and $\mathcal{L}(f)$. Here, the Laplace transform of f (i.e. $\mathcal{L}(f)$) is known, but not necessarily f .

f is replaced by the inverse Laplace transform (defined in (4.35)) of $\mathcal{L}(f)$ in (4.57)

$$x(t) = \frac{1}{2\pi i} \int_{s=\sigma-i\infty}^{\sigma+i\infty} \mathcal{L}(f)(s) \int_{u=0}^t e^{s(t-u)}g(u)duds. \quad (4.59)$$

Note that the inner integral

$$y(t) = \int_{u=0}^t e^{s(t-u)}g(u)du \quad (4.60)$$

is solution of the differential equation (4.51). Using the solution of the IRK (4.50) in (4.57) results in

$$x_{j+1} \approx \frac{1}{2\pi i} \int_{s=\sigma-i\infty}^{\sigma+i\infty} \mathcal{L}(f)(s)\mathbf{b}^\top\mathbf{A}^{-1}\mathbf{Y}_j ds. \quad (4.61)$$

Then using (4.54), its Z-transform is

$$\mathcal{Z}((x_{j+1})_j)(z) \approx \frac{1}{2\pi i} \int_{s=\sigma-i\infty}^{\sigma+i\infty} \mathcal{L}(f)(s)\mathbf{b}^\top\mathbf{A}^{-1}\mathcal{Z}(\mathbf{Y})(z)ds \quad (4.62a)$$

$$\approx \mathbf{b}^\top\mathbf{A}^{-1} \left(\frac{1}{2\pi i} \int_{s=\sigma-i\infty}^{\sigma+i\infty} \mathcal{L}(f)(s) \cdot (\mathbf{s}(z) - s\mathbf{I}_p)^{-1} ds \right) \mathcal{Z}(\mathbf{g})(z). \quad (4.62b)$$

The next step is to simplify this integral. Assuming that a function $X(s)$ has a power series expansion

$$X(s) = \sum_{n=0}^{+\infty} \frac{X^{(n)}(0)}{n!} s^n, \quad (4.63)$$

then, this function applied on a matrix \mathbf{s} is noted $X(\mathbf{s})$ and defined as

$$X(\mathbf{s}) = \sum_{n=0}^{+\infty} \frac{X^{(n)}(0)}{n!} \mathbf{s}^n. \quad (4.64)$$

With this notation the integral simplifies to [49]

$$\frac{1}{2\pi i} \int_{s=\sigma-i\infty}^{\sigma+i\infty} \mathcal{L}(f)(s) \cdot (\mathbf{s}(z) - s\mathbf{I}_p)^{-1} ds = \mathcal{L}(f)(\mathbf{s}(z)) \quad (4.65)$$

which can be obtained by rewriting the matrix inverse as a power series, then building a closed contour for the integral by adding an half circle at infinity that enclose the right hand plane, and finally using the Cauchy integral formula for derivatives. Therefore,

$$\mathcal{Z}((x_{i+1})_i)(z) \approx \mathbf{b}^\top \mathbf{A}^{-1} \mathcal{L}(f)(\mathbf{s}(z)) \cdot \mathcal{Z}(\mathbf{g})(z). \quad (4.66)$$

Inverse Z-transforming (see (4.37)) the previous equation results in the convolution quadrature

$$x_{i+1} \approx (\mathbf{w} * \mathbf{g})_i = \sum_{j=0}^i \mathbf{w}_{i-j} \mathbf{g}_j \quad (4.67)$$

where $[\mathbf{g}_j]_k = g(t_j + \mathbf{c}_k \Delta t)$ and the weights $(\mathbf{w}_j)_j$ are a sequence of row vectors in $\mathbb{R}^{1 \times p}$

$$\mathbf{w}_j = \mathbf{b}^\top \mathbf{A}^{-1} \mathcal{Z}^{-1}(z \mapsto \mathcal{L}(f)(\mathbf{s}(z)))_j \quad (4.68)$$

It can be read from (4.68) that the weights are actually the inverse Z-transform of the kernel in the Laplace domain evaluated at matrix valued frequencies. Also, the weights are vectors that contain the interpolants for the p stages of the IRK method. Note that if the IRK method is L-stable i.e. $\mathbf{b}^\top \mathbf{A}^{-1} = (0, \dots, 0, 1)$, then \mathbf{w}_j is actually the last row of the matrix $\mathcal{Z}^{-1}(\mathcal{L}(f)(\mathbf{s}(z)))_j$.

4.3.4 Discretization of time domain integral equations

In this section, the differentiated PEC-EFIE (4.16) is used as an example to discretize a time domain integral equation using the convolution quadrature. Again, the equation to discretize is

$$\eta \frac{\partial}{\partial t} \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) = -\hat{\mathbf{n}}(\mathbf{r}) \times \frac{\partial}{\partial t} \mathbf{E}^{\text{inc}}(\mathbf{r}, t). \quad (4.69)$$

The spacial discretization is achieved using a Galerkin strategy with RWG basis functions (\mathbf{f}_n) , similarly to the time testing scheme. The current is discretized as

$$\mathbf{J}_s(\mathbf{r}, t) \approx \sum_{n=1}^{N_s} [\mathbf{J}(t)]_n \mathbf{f}_n(\mathbf{r}), \quad (4.70)$$

4. Introduction to the time domain

and the system is

$$(\mathbf{Z} * \mathbf{J})(t) = \mathbf{E}(t) \quad (4.71)$$

where for all $m, n \in [1, N_s]$,

$$[\mathbf{Z}(t)]_{mn} = \left\langle \hat{\mathbf{n}} \times \mathbf{f}_m, \eta \frac{\partial}{\partial t} \mathcal{T}(\mathbf{f}_n \delta)(t) \right\rangle \quad (4.72a)$$

$$= \eta \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial}{\partial t} \mathcal{T}(\mathbf{f}_n \delta)(\mathbf{r}, t) \, dS \quad (4.72b)$$

$$[\mathbf{E}(t)]_m = \left\langle \mathbf{f}_m, \frac{\partial \mathbf{E}^{\text{inc}}}{\partial t}(t) \right\rangle \quad (4.72c)$$

$$= - \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial \mathbf{E}^{\text{inc}}}{\partial t}(\mathbf{r}, t) \, dS \quad (4.72d)$$

where the Dirac delta $\delta(t)$ is represented in the figure 4.2. Applying the convolution quadrature (4.67) on (4.71) results in

$$[\mathbf{E}(t_{i+1})]_m = \sum_{n=1}^{N_s} ([\mathbf{Z}]_{mn} * [\mathbf{J}]_n)(t_{i+1}) \quad (4.73a)$$

$$= \mathbf{b}^\top \mathbf{A}^{-1} \sum_{j=0}^i \sum_{n=1}^{N_s} \mathcal{Z}^{-1}(z \mapsto \mathcal{L}([\mathbf{Z}]_{mn})(\mathbf{s}(z)))_{i-j} \mathbf{g}_{n,j} \quad (4.73b)$$

for each $m \in [1, N_s]$ and where the elements of $\mathbf{g}_{n,j} \in \mathbb{R}^p$ are $[\mathbf{g}_{n,j}]_k = [\mathbf{J}(t_j + \mathbf{c}_k \Delta t)]_n$ ($k \in [1, p]$). The Laplace domain MoM matrix $\mathcal{L}(\mathbf{Z})(s) \in \mathbb{C}^{N_s \times N_s}$ evaluated at a matrix valued frequency $\mathbf{s}(z) \in \mathbb{C}^{p \times p}$ should be interpreted as a block matrix in $\mathbb{C}^{pN_s \times pN_s}$ such that for all $k, l \in [1, p]$ and $m, n \in [1, N_s]$

$$[\mathcal{L}(\mathbf{Z})(\mathbf{s}(z))]_{k+(m-1)p, l+(n-1)p} = [\mathcal{L}(\mathbf{Z}_{mn})(\mathbf{s}(z))]_{k,l}. \quad (4.74)$$

According to this, the $\mathbf{g}_{n,j} \in \mathbb{R}^p$ in (4.73b) are stored in a sequence of vectors $(\mathbf{J}_j)_{j \in \mathbb{N}}$ in \mathbb{R}^{pN_s} and for all $k \in [1, p]$, $m \in [1, N_s]$ and $j \in \mathbb{N}$

$$[\mathbf{J}_j]_{k+(n-1)p} = [\mathbf{J}(t_j + \mathbf{c}_k \Delta t)]_n = [\mathbf{g}_{n,j}]_k \quad (4.75a)$$

$$\mathbf{J}_s(\mathbf{r}, t_j + \mathbf{c}_k \Delta t) \approx \sum_{n=1}^{N_s} [\mathbf{J}_j]_{k+(n-1)p} \mathbf{f}_n(\mathbf{r}). \quad (4.75b)$$

To convert (4.73b) into a Marching-On-in-Time (MOT) scheme, a sequence of excitation vectors in \mathbb{R}^{pN_s} , noted $(\mathbf{E}_i)_{i \in \mathbb{N}}$, is built as

$$[\mathbf{E}_i]_{k+(m-1)p} = [\mathbf{E}(t_i + \mathbf{c}_k \Delta t)]_m \quad (4.76a)$$

$$= - \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \frac{\partial \mathbf{E}^{\text{inc}}}{\partial t}(\mathbf{r}, t_i + \mathbf{c}_k \Delta t) \, dS. \quad (4.76b)$$

Then, assuming the L-stability of the Runge-Kutta method, (4.73b) becomes for all $i \in \mathbb{N}$

$$\sum_{j=0}^i \mathbf{z}_{i-j} \mathbf{J}_j = \mathbf{E}_i \quad (4.77)$$

or equivalently

$$\sum_{j=0}^i \mathbf{z}_j \mathbf{J}_{i-j} = \mathbf{E}_i \quad (4.78)$$

where the sequence of interaction matrices $(\mathbf{z}_j)_j$ is defined as

$$\mathbf{z}_j = \mathcal{Z}^{-1}(z \mapsto \mathcal{L}(t \mapsto \mathbf{Z}(t))(\mathbf{s}(z)))_j. \quad (4.79)$$

The sequence $(\mathbf{z}_j)_j$ must not be confused with $\mathbf{Z}(t)$ (4.72a). The sequence $(\mathbf{z}_j)_j$ is the inverse \mathcal{Z} -transform of $\mathcal{L}(t \mapsto \mathbf{Z}(t))$ evaluated at the matrix valued frequency $\mathbf{s}(z)$. (4.78) is rewritten to have a MOT scheme as

$$\mathbf{z}_0 \mathbf{J}_i = \mathbf{E}_i - \sum_{j=1}^i \mathbf{z}_j \mathbf{J}_{i-j} \quad (4.80)$$

Note that there is a minor difference with the MOT (4.32) obtained from the time testing in section 4.2 as a result of the indexing convention:

- With the time testing, \mathbf{J}_0 scales the current for the pulse $p(t - t_0)$ (4.23) which is non-zero between t_{-1} and t_0 . Therefore $\mathbf{J}_0 = \mathbf{0}$ from the assumption that $\mathbf{J}_s(\mathbf{r}, t) = \mathbf{0}$ for all $t < t_0$. So \mathbf{J}_0 is not used in the MOT (4.32).
- With the convolution quadrature, \mathbf{J}_0 samples the current for t between t_0 and t_1 (4.75a). Therefore, \mathbf{J}_0 is non zero and it is used in the MOT (4.80) (and $\mathbf{J}_i = \mathbf{0}$ for all $i < 0$).

In the section 4.2, the time testing procedure resulted in a sequence of matrices $(\mathbf{z}_j)_j$ that are identically zero after some rank (roughly $\max_{\mathbf{r}, \mathbf{r}' \in \Gamma} \frac{|\mathbf{r} - \mathbf{r}'|}{c\Delta t}$). But, this is not the case for the sequence $(\mathbf{z}_j)_j$ stemming from the convolution quadrature. However,

$$\|\mathbf{z}_j\| \xrightarrow{j \rightarrow +\infty} 0 \quad (4.81)$$

exponentially. So in practice, there is a rank after which the norms of all the matrix in the sequence go below the machine precision, thus the convolution in the MOT (4.80) can be truncated to N_{conv} (not necessarily the same as in (4.33)). So,

$$\mathbf{z}_0 \mathbf{J}_i = \mathbf{E}_i - \sum_{j=1}^{N_{\text{conv}}} \mathbf{z}_j \mathbf{J}_{i-j}. \quad (4.82)$$

4.3.5 Computation of the matrix elements

4.3.5.1 Operators in the Laplace domain

This section details the computation of $\mathcal{L}(t \mapsto \mathbf{Z}(t))(\mathbf{s}(z))$ in (4.74). The analogues of the operators (4.15) in the Laplace domain are

$$\mathcal{T}\mathbf{f}(\mathbf{r}, s) = -\frac{s}{c}\mathcal{T}_s\mathbf{f}(\mathbf{r}, s) + \frac{c}{s}\mathcal{T}_h\mathbf{f}(\mathbf{r}, s) \quad (4.83a)$$

$$\mathcal{T}_s\mathbf{f}(\mathbf{r}, s) = \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-\frac{s}{c}|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \mathbf{f}(\mathbf{r}', s) dS' \quad (4.83b)$$

$$\mathcal{T}_h\mathbf{f}(\mathbf{r}, s) = \hat{\mathbf{n}}(\mathbf{r}) \times \nabla \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-\frac{s}{c}|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \nabla' \cdot \mathbf{f}(\mathbf{r}', s) dS' \quad (4.83c)$$

$$\mathcal{K}\mathbf{f}(\mathbf{r}, s) = \hat{\mathbf{n}}(\mathbf{r}) \times \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla \left(\frac{e^{-\frac{s}{c}|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \right) \times \mathbf{f}(\mathbf{r}', s) dS'. \quad (4.83d)$$

In fact, these operators can be simply obtained from the expression of the operators in the frequency domain with the substitution of ik by s/c in (1.68a) and (1.69).

In the Laplace domain, the PEC-EFIE reads

$$\eta\mathcal{T}\mathbf{J}_s(\mathbf{r}, s) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, s). \quad (4.84)$$

In fact, the EFIE is usually time differentiated to be solved. It enables getting rid of the $1/s$ factor in front of \mathcal{T}_h in (4.83a) that corresponds to a time integration. Eventually, this allows the convolution kernel on the RHS of the MOT (4.82) to be truncated to N_{conv} terms as the norms of the matrices $\|\mathbf{Z}_j\|$ decay exponentially, which is not the case for the non-differentiated EFIE. Note that in the case of testing with temporal basis functions, it enables to get interaction matrices that are exactly $\mathbf{0}$ as explained before the equation (4.33). The differentiated PEC-EFIE in the Laplace domain is

$$s\eta\mathcal{T}\mathbf{J}_s(\mathbf{r}, s) = -s\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, s). \quad (4.85)$$

A Galerkin discretization in space with RWG basis functions (1.90) results in the linear system

$$\mathbf{Z}(s)\mathbf{J}(s) = \mathbf{E}(s) \quad (4.86)$$

where for all $m, n \in [1, N_s]$

$$\mathbf{J}_s(\mathbf{r}, s) \approx \sum_{n=1}^{N_s} [\mathbf{J}(s)]_n \mathbf{f}_n(\mathbf{r}) \quad (4.87a)$$

$$[\mathbf{Z}(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, s\eta\mathcal{T}\mathbf{f}_n(s) \rangle \quad (4.87b)$$

$$[\mathbf{E}(s)]_m = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, -s\hat{\mathbf{n}} \times \mathbf{E}^{\text{inc}}(s) \rangle. \quad (4.87c)$$

The relation between the Laplace domain $\mathbf{Z}(s)$ in (4.87b) and the time domain $\mathbf{Z}(t)$ in (4.72a) is $\mathbf{Z}(s) = \mathcal{L}(t \mapsto \mathbf{Z}(t))(s)$ as expected.

4.3.5.2 Computation of the matrix elements in the \mathbf{Z} -domain

$\mathbf{Z}(s)$ is evaluated in a matrix valued argument $\mathbf{s}(z)$. In practice, the matrix $\mathbf{s}(z)$ is diagonalized as

$$\mathbf{s}(z) = \mathbf{M}(z)\mathbf{D}(z)\mathbf{M}(z)^{-1} \quad (4.88)$$

where \mathbf{D} is a diagonal matrix made of the p eigenvalues of \mathbf{s} (so the eigenvalues are \mathbf{D}_{kk} for $k \in [1, p]$), and \mathbf{M} is the matrix that contains the p eigenvectors in each column. For $m, n \in [1, N_s]$, note the matrix entry $X(s) = [\mathbf{Z}(s)]_{mn}$. Then, $X(\mathbf{s}) = \mathbf{M}X(\mathbf{D})\mathbf{M}^{-1}$ where $X(\mathbf{D})$ is a diagonal matrix whose k -th entry is $X(\mathbf{D}_{kk})$. Applying this process on each element of $\mathbf{Z}(s)$ results in

$$\mathbf{Z}(\mathbf{s}) = \tilde{\mathbf{M}}\mathbf{Z}(\mathbf{D})\tilde{\mathbf{M}}^{-1} \quad (4.89)$$

where for $m, n \in [1, N_s]$ and $k \in [1, p]$

$$[\mathbf{Z}(\mathbf{D})]_{k+p(m-1), k+p(n-1)} = [\mathbf{Z}(\mathbf{D}_{kk})]_{mn} \quad (4.90)$$

and $\tilde{\mathbf{M}}$ is a block diagonal matrix with N_s blocks equal to \mathbf{M}

$$\tilde{\mathbf{M}} = \begin{pmatrix} \mathbf{M} & & & \\ & \mathbf{M} & & (\mathbf{0}) \\ & & \ddots & \\ & (\mathbf{0}) & & \mathbf{M} \end{pmatrix}. \quad (4.91)$$

From (4.89) and (4.90) it is clear that the matrix valued frequency dependent $\mathbf{Z}(\mathbf{s}(z))$ can be obtained from p computations of the scalar valued frequency dependent $\mathbf{Z}(\mathbf{D}_{kk}(z))$. The computation of these matrices in the Laplace domain is actually the same as the matrices in the Fourier domain that is described in the section 1.3.4 and where ik has been replaced by s/c .

4.3.5.3 Kronecker product

For all $\mathbf{X} \in \mathbb{C}^{N_s \times N_s}$ and $\mathbf{Y} \in \mathbb{C}^{p \times p}$, the Kronecker product $\mathbf{X} \otimes \mathbf{Y}$ is defined as

$$[\mathbf{X} \otimes \mathbf{Y}]_{k+p(m-1), l+p(n-1)} = [\mathbf{X}]_{mn} [\mathbf{Y}]_{kl} \quad (m, n \in [1, N_s] \text{ and } k, l \in [1, p]). \quad (4.92)$$

For example using this notation, $\tilde{\mathbf{M}}$ in (4.91) can be conveniently rewritten as $\tilde{\mathbf{M}} = \mathbf{I} \otimes \mathbf{M}$ where \mathbf{I} is the $N_s \times N_s$ identity matrix.

In fact, in the next chapters on the time domain, an extensive use of the Kronecker product \otimes is done. This product is used in two cases:

- With the identity matrix $\mathbf{I} \in \mathbb{R}^{N_s \times N_s}$ and another matrix $\mathbf{Y} \in \mathbb{C}^{p \times p}$. The result is noted $\tilde{\mathbf{Y}} = \mathbf{I} \otimes \mathbf{Y}$.
- With a matrix $\mathbf{X} \in \mathbb{C}^{N_s \times N_s}$ and the identity matrix $\mathbf{I}_p \in \mathbb{R}^{p \times p}$. The result is noted $\tilde{\mathbf{X}} = \mathbf{X} \otimes \mathbf{I}_p$.

The tilde is used to mark the augmented matrix. Similarly to what has been done in the previous section with $\tilde{\mathbf{M}}$, the Kronecker product with an identity enables to perform a block-wise multiplication. This is convenient in the case of IRK methods that yield quantities with p stages after the substitution of the scalar s by a matrix \mathbf{s} .

For example, consider a constant (relative to s) matrix $\mathbf{X} \in \mathbb{C}^{N_s \times N_s}$. Then, $\mathbf{Z}(s) = \mathbf{X}\mathbf{T}(s)$ becomes $\mathbf{Z}(\mathbf{s}) = \tilde{\mathbf{X}}\mathbf{T}(\mathbf{s})$ after substitution, where $\tilde{\mathbf{X}} = \mathbf{X} \otimes \mathbf{I}_p$. Intuitively, each entry of \mathbf{X} is independent of s so they can be thought as being multiplied by s^0 . Therefore, substituting s by \mathbf{s} transforms the factor s^0 by $\mathbf{s}^0 = \mathbf{I}_p$.

The second example is the multiplication by a scalar function $Y(s)$. Then, $\mathbf{Z}(s) = Y(s)\mathbf{T}(s)$ becomes $\mathbf{Z}(\mathbf{s}) = \tilde{\mathbf{Y}}(\mathbf{s})\mathbf{T}(\mathbf{s})$ after the substitution, where $\tilde{\mathbf{Y}}(\mathbf{s}) = \mathbf{I} \otimes Y(\mathbf{s})$. The matrix $Y(\mathbf{s}) \in \mathbb{C}^{p \times p}$ is defined from the scalar $Y(s) \in \mathbb{C}$ using (4.64). Again, the multiplication of $Y(s)$ by a matrix (or vector) can be thought as a multiplication with the diagonal matrix with $Y(s)$ on all the diagonal. After the substitution of s by \mathbf{s} , it transforms into a block diagonal matrix with $Y(\mathbf{s})$ on all the diagonal i.e. $\mathbf{I} \otimes Y(\mathbf{s})$.

In addition, note that $\mathbf{s} = \mathbf{M}\mathbf{D}\mathbf{M}^{-1}$ is transformed block-wise into $\tilde{\mathbf{s}} = \tilde{\mathbf{M}}\tilde{\mathbf{D}}\tilde{\mathbf{M}}^{-1}$ where $\tilde{\mathbf{s}} = \mathbf{I} \otimes \mathbf{s}$, $\tilde{\mathbf{M}} = \mathbf{I} \otimes \mathbf{M}$, $\tilde{\mathbf{D}} = \mathbf{I} \otimes \mathbf{D}$ and $\tilde{\mathbf{M}}^{-1} = \mathbf{I} \otimes (\mathbf{M}^{-1})$ (note that the inverse is block-wise). This is applied to $\mathbf{Z}(s) = \eta s \mathbf{T}(s)$ in (4.87b) which results in $\mathbf{Z}(\mathbf{s}) = \eta \tilde{\mathbf{s}} \mathbf{T}(\mathbf{s}) = \eta \tilde{\mathbf{M}} \tilde{\mathbf{D}} \mathbf{T}(\mathbf{s}) \tilde{\mathbf{M}}^{-1}$.

4.4 DC instability

Similarly to the frequency domain integral equations, the time domain integral equations also have problems similar to the low frequency breakdown, dense mesh breakdown and numerical cancellations. In the time domain, the low frequency breakdown is the so-called large time step breakdown. However, there is also the so-called DC instability that is specific to the time domain. In this section, the origin of the DC instability is explained. In fact, both the testing with temporal basis functions (section 4.2) and the convolution quadrature (section 4.3) time discretizations have these problems.

Consider the PEC-EFIE

$$\eta \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, t). \quad (4.93)$$

Let \mathbf{J}_s^{cs} be a constant (in time) solenoidal current i.e. $\frac{\partial \mathbf{J}_s^{\text{cs}}}{\partial t} = \mathbf{0}$ and $\nabla \cdot \mathbf{J}_s^{\text{cs}} = 0$. Then, it is clear from the definition of \mathcal{T} (4.15a) that $\mathcal{T} \mathbf{J}_s^{\text{cs}} = \mathbf{0}$. In other words, if \mathbf{J}_s is solution of the PEC-EFIE, then $\mathbf{J}_s + \mathbf{J}_s^{\text{cs}}$ is also solution. To enforce the uniqueness and the causality of the solution, the current \mathbf{J}_s is assumed to be zero when $t < t_0$ so that the only possible value for \mathbf{J}_s^{cs} is zero.

Note that in the case of the differentiated PEC-EFIE, also the linear (in time) solenoidal currents are in the null space of the integral equation operator i.e. if

$$\frac{\partial^2 \mathbf{J}_s^{\text{ls}}}{\partial t^2} = \mathbf{0} \text{ and } \nabla \cdot \mathbf{J}_s^{\text{ls}} = 0 \text{ then } \frac{\partial}{\partial t} \mathcal{T} \mathbf{J}_s^{\text{ls}} = \mathbf{0}.$$

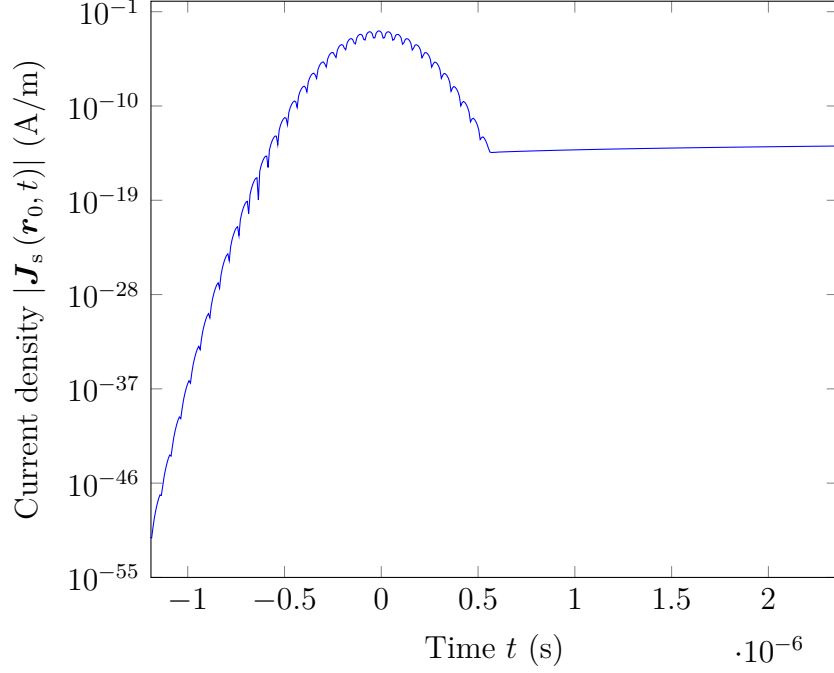


Figure 4.3: Surface current density on the unit sphere showing the DC instability at $\mathbf{r}_0 = (0 \text{ m}, 0.93 \text{ m}, 0.36 \text{ m})$

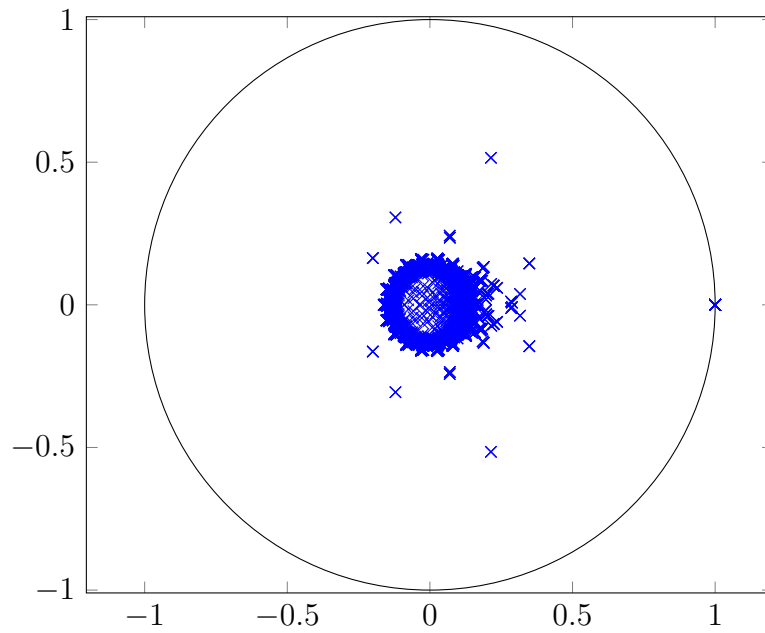
Unfortunately, even if these constant or linear in time solenoidal currents are enforced to be zero in the continuous integral equation by imposing $\mathbf{J}_s(\mathbf{r}, t) = \mathbf{0}$ for all $t < t_0$, they still plague the discrete solution because they are in the null space of the integral equation operator. Because of numerical errors (approximations and floating point truncation errors), these static currents appear in the late time steps as a slowly (yet exponentially) growing non-physical solution of the discrete equation. This current has been plotted in the figure 4.3 for a unit sphere with the parameters $N_s = 120$, $N_t = 500$, $N_{\text{conv}} = 17$, $\Delta t = 7.16 \text{ ns}$, IRK method (4.45) and an incident modulated Gaussian plane wave

$$\mathbf{E}^{\text{inc}}(\mathbf{r}, t) = \hat{\mathbf{p}} \exp\left(-\frac{\tau_d(\mathbf{r}, t)^2}{2\tau_{\text{ch}}^2}\right) \cos(2\pi f_0 \tau_d(\mathbf{r}, t)) E_0 \quad (4.94)$$

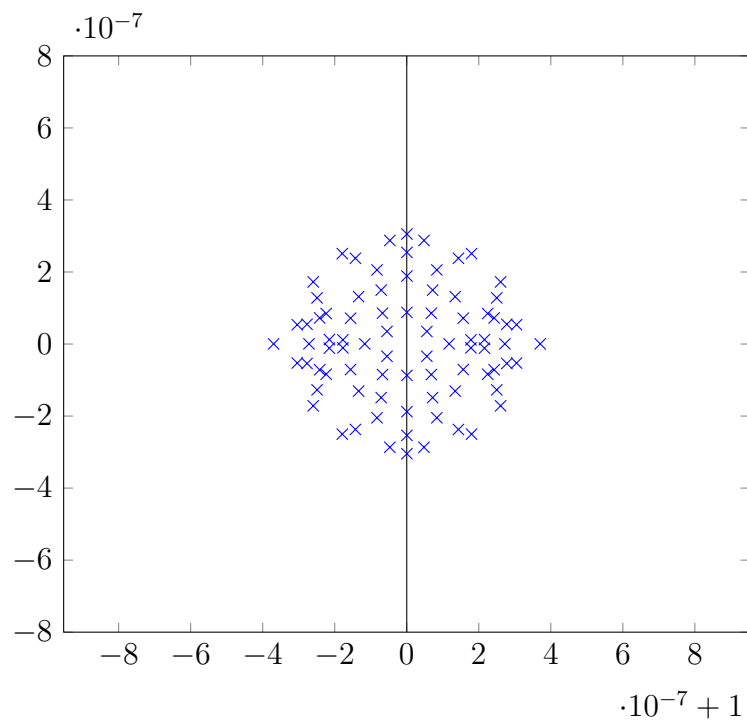
where $\hat{\mathbf{k}} = -\hat{\mathbf{z}}$, $\tau_d(\mathbf{r}, t) = t - \frac{1}{c} \hat{\mathbf{k}} \cdot \mathbf{r}$, $\hat{\mathbf{p}} = \hat{\mathbf{x}}$, $E_0 = 1 \text{ V/m}$, $f_0 = 10 \text{ MHz}$ and $\tau_{\text{ch}} = 79.6 \text{ ns}$.

The DC instability can be illustrated by studying the MOT for the homogeneous system i.e. without excitation ($\mathbf{E}_i = \mathbf{0}$ for all i)

$$\mathbf{Z}_0 \mathbf{J}_i = - \sum_{j=1}^{N_{\text{conv}}} \mathbf{Z}_j \mathbf{J}_{i-j}. \quad (4.95)$$



(a) Polynomial eigenvalues with a cluster in 1



(b) Cluster of polynomial eigenvalues in 1

Figure 4.4: Polynomial eigenvalues in the complex plane

Define the block vector \mathbf{J}_i containing N_{conv} successive discrete solutions

$$\mathbf{J}_i = \begin{pmatrix} \mathbf{J}_i \\ \mathbf{J}_{i-1} \\ \vdots \\ \mathbf{J}_{i+1-N_{\text{conv}}} \end{pmatrix} \quad (4.96)$$

and the block matrix \mathbf{Z}_C called companion matrix [50]

$$\mathbf{Z}_C = \begin{pmatrix} -\mathbf{Z}_0^{-1}\mathbf{Z}_1 & -\mathbf{Z}_0^{-1}\mathbf{Z}_2 & \cdots & -\mathbf{Z}_0^{-1}\mathbf{Z}_{N_{\text{conv}}-1} & -\mathbf{Z}_0^{-1}\mathbf{Z}_{N_{\text{conv}}} \\ \mathbf{I} & & & & \\ & \mathbf{I} & & & \\ & & \ddots & & \\ & & & \mathbf{I} & \\ & & & & \mathbf{I} \end{pmatrix}. \quad (4.97)$$

They verify the recurrence relation

$$\mathbf{J}_i = \mathbf{Z}_C \mathbf{J}_{i-1}. \quad (4.98)$$

So by studying the eigenvalues of \mathbf{Z}_C , also called polynomial eigenvalues, it is possible to understand the behaviour of the solution \mathbf{J}_i when $i \rightarrow +\infty$. If all the eigenvalues (noted λ_n) of \mathbf{Z}_C , are in the interior of the unit circle (i.e. $|\lambda_n| < 1$ for all n), then the system is stable, since any numerical perturbation in the solution vanishes exponentially. However, if an eigenvalue is outside of the unit circle (i.e. there is a n such that $|\lambda_n| > 1$), then the system is unstable and any numerical error in the solution grows exponentially.

This is illustrated in the figure 4.4 where the polynomial eigenvalues for the differentiated PEC-EFIE have been plotted in the complex plane (with the same parameters as before). The eigenvalues associated to the constant and linear in time solenoidal current are clustered around 1 which results in the DC instability. They are not exactly on 1 because of numerical errors and approximations in the discretized system and some of them lie outside of the unit circle.

Chapter 5

Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods

To simulate the transient scattering of a Perfect Electric Conductor (PEC) in the time domain, the PEC-EFIE and its differentiated version are widely used. In this chapter, the discretization of the PEC-EFIE is obtained with a Galerkin strategy using RWG basis functions in space, and a convolution quadrature using Implicit Runge-Kutta (IRK) methods in time. Unfortunately, the resulting system has two problems. First, the linear system suffers from a large time step breakdown i.e. it becomes ill-conditioned when the time step increases. And second, there is a DC instability i.e. the formulation allows the existence of spurious static solenoidal currents in the solution. In this chapter, these two problems are solved by leveraging the quasi-Helmholtz projectors. Besides, it is shown how the projectors can be used with the IRK based convolution quadrature.

5.1 Introduction

The Perfect Electric Conductor - Electric Field Integral Equation (PEC-EFIE) and more particularly the differentiated PEC-EFIE are widely used to simulate a PEC scatterer in the time domain [51]. The numerical solution of time domain integral equations has received many research contributions to reduce the computational cost with the Plane Wave Time Domain (PWTD) [44, 45] algorithm or the Hierarchical-FFT (HIL-FFT) [52]. The stability of the equations has been addressed using loop-tree decompositions [53, 54], Calderón preconditioning [55], quasi-Helmholtz projectors [56] and Combined Field Integral Equations (CFIE) [57, 58]. Finally, the accuracy of the equations can be improved using higher order spacial basis functions [59, 60], better temporal basis functions [61–63] and exact evaluations [64, 65].

A vast majority of formulations discretizes the equation in space and time which

results in a fully discretized system that can be solved by Marching-On-in-Time (MOT), although it exists other approaches such as marching-on-in-order [66]. The space discretization is commonly achieved using a space Galerkin testing with the set of basis functions that spans the space of surface current densities. The system is then discretized in time, amongst others, with point testing [51], space-time Galerkin [67] or Runge-Kutta convolution quadrature [47, 48]. Implicit Runge-Kutta (IRK) methods are known to have a good stability in solving ordinary differential equation. These good properties are inherited when the IRK methods are used to discretize convolution quadratures with matrices stemming from a Boundary Element Method (BEM) discretization. Also, it only requires the Laplace domain expression of the convolution kernel, which has the practical advantage of getting a time domain solver from a frequency domain code. This opposes to the space-time Galerkin that requires more complicated quadrature schemes for the computation of the interaction matrices elements.

The differentiated PEC-CFIE [68] and the differentiated PEC-EFIE [69] can be solved using an IRK convolution quadrature. In particular, the PEC-EFIE has two problems:

- It suffers from the large time step breakdown. It causes the condition number of the matrix in the MOT system to grow quadratically with time step [55]. This is in fact the time domain counterpart of the low frequency breakdown.
- It suffers from the DC instability. The equation allows the existence of static solenoidal currents in the solution [50]. In practice it yields a wrong current in the late time steps.

In [56], these two problems were solved for the space-time Galerkin discretization using quasi-Helmholtz projectors. By differentiating or integrating the appropriate Helmholtz components, the resulting system is free from both the large time step breakdown and the DC instability. Furthermore, the projectors do not require the expensive detection of the global loops in case of multiply connected geometry.

In this chapter, it is shown how the quasi-Helmholtz projectors can be used in the context of IRK convolution quadrature discretizations. Using the convolution quadrature on the MoM system preconditioned in the Laplace domain results in a regularized MOT scheme that does not suffer from the large time step breakdown nor the DC instability.

The chapter is subdivided in five sections. First, the groundwork of this chapter that is the PEC-EFIE discretized with IRK convolution quadrature is briefly recalled. Then the two problems, i.e. the large time step breakdown and the DC instability, are analyzed in the context of a time discretization based on an IRK convolution quadrature. In the next section, the quasi-Helmholtz projectors are used to regularize the PEC-EFIE. Then, an analysis of the computational complexity of the solver is done to show that the quasi-Helmholtz projectors are compatible with the IRK convolution quadrature. And finally, some numerical results are presented to illustrate the work.

Preliminary results on this work were presented at a conference [70].

5.2 Background and notations

5.2.1 Perfect Electric Conductor - Electric Field Integral Equation

5.2.1.1 PEC-EFIE in the time domain

A Perfect Electric Conducting (PEC) scatterer whose boundary is Γ is placed in a background medium that has a permittivity ε , a permeability μ and characteristic impedance $\eta = \sqrt{\mu/\varepsilon}$. An incident electric field \mathbf{E}^{inc} induces a surface electric current density \mathbf{J}_s on Γ . The current \mathbf{J}_s is solution of the time domain PEC-EFIE (4.14a)

$$\eta \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, t) \quad (5.1)$$

where the operator \mathcal{T} is defined in (4.15a) and $\hat{\mathbf{n}}$ is the vector normal to Γ . The time differentiated PEC-EFIE is also considered as it is often used to avoid the time integration of the hypersingular part of \mathcal{T}

$$\eta \frac{\partial}{\partial t} \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) = -\hat{\mathbf{n}}(\mathbf{r}) \times \frac{\partial \mathbf{E}^{\text{inc}}}{\partial t}(\mathbf{r}, t). \quad (5.2)$$

In order to enforce the causality and the uniqueness of the solution, all the fields are assumed to vanish in the neighbourhood of Γ when $t < t_0$ (see section 4.4).

5.2.1.2 PEC-EFIE in the Laplace domain

The PEC-EFIE in the Laplace domain reads

$$\eta \mathcal{T} \mathbf{J}_s(\mathbf{r}, s) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, s) \quad (5.3)$$

where the Laplace domain operator \mathcal{T} (4.83a) is

$$\mathcal{T} \mathbf{J}_s(\mathbf{r}, s) = -\frac{s}{c} \mathcal{T}_s \mathbf{J}_s(\mathbf{r}, s) + \frac{c}{s} \mathcal{T}_h \mathbf{J}_s(\mathbf{r}, s) \quad (5.4a)$$

$$\mathcal{T}_s \mathbf{J}_s(\mathbf{r}, s) = \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-\frac{s}{c}|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \mathbf{J}_s(\mathbf{r}', s) dS' \quad (5.4b)$$

$$\mathcal{T}_h \mathbf{J}_s(\mathbf{r}, s) = \hat{\mathbf{n}}(\mathbf{r}) \times \nabla \iint_{\mathbf{r}' \in \Gamma} \frac{e^{-\frac{s}{c}|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \nabla' \cdot \mathbf{J}_s(\mathbf{r}', s) dS' \quad (5.4c)$$

with $c = 1/\sqrt{\mu\varepsilon}$ that denotes the speed of light in the background medium. The time derivative is transformed to a factor s in the Laplace domain, so the time differentiated PEC-EFIE transformed to the Laplace domain reads

$$\eta s \mathcal{T} \mathbf{J}_s(\mathbf{r}, s) = -s \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, s). \quad (5.5)$$

5.2.2 Space Galerkin discretization

Γ is approximated by a mesh containing N_s edges, N_v vertices and N_f triangular faces. A set of N_s Rao-Wilton-Glisson [3] (RWG) basis functions $(\mathbf{f}_n)_{n=1}^{N_s}$ (1.90) is built on Γ . In the Laplace domain, the current is expanded in the RWG basis as

$$\mathbf{J}_s(\mathbf{r}, s) \approx \sum_{m=1}^{N_s} [\mathbf{J}(s)]_m \mathbf{f}_m(\mathbf{r}) \quad (5.6)$$

where $\mathbf{J}(s) \in \mathbb{C}^{N_s}$ is the array of coefficients of the expansion. A Galerkin testing procedure on (5.3) results in the following system in the Laplace domain

$$\mathbf{Z}(s)\mathbf{J}(s) = \mathbf{E}(s) \quad (5.7)$$

with

$$\mathbf{Z}(s) = \eta \left(-\frac{s}{c} \mathbf{T}_s(s) + \frac{c}{s} \mathbf{T}_h(s) \right) \quad (5.8)$$

$$[\mathbf{T}_s(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_s \mathbf{f}_n(s) \rangle \quad (5.9)$$

$$[\mathbf{T}_h(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_h \mathbf{f}_n(s) \rangle \quad (5.10)$$

$$[\mathbf{E}(s)]_m = -\langle \mathbf{f}_m, \mathbf{E}^{\text{inc}}(s) \rangle. \quad (5.11)$$

The computation of the above elements is detailed in the appendix A.

The system corresponding to the time differentiated PEC-EFIE transformed to the Laplace domain is

$$s\mathbf{Z}(s)\mathbf{J}(s) = s\mathbf{E}(s). \quad (5.12)$$

5.2.3 Time discretization with IRK convolution quadrature

The time discretization is achieved with a convolution quadrature using IRK methods. As explained in the section 4.3, the time domain system is transformed to the Laplace domain, then discretized to the Z-domain using IRK methods, and finally transformed to the discrete time domain with an inverse Z-transform.

The solution is assumed to be $\mathbf{0}$ for all $t < t_0$. Then, it is computed for each $t_i = t_0 + i\Delta t$ where Δt is the time step and $i \leq N_t$. The system in the Laplace domain (5.7) is discretized by expressing the Laplace variable s as a function of z (the Z-domain variable). This is achieved using Implicit Runge-Kutta methods. It formally corresponds to the substitution of the Laplace variable s by a matrix valued $\mathbf{s}(z)$ (4.55)

$$\mathbf{s}(z) = \frac{1}{\Delta t} \left(\mathbf{A} + \frac{\mathbf{1b}^T}{z-1} \right)^{-1} \quad (5.13)$$

where $\mathbf{A} \in \mathbb{R}^{p \times p}$ and $\mathbf{b} \in \mathbb{R}^p$ (p is the number of stages in the IRK method) are given by the Butcher tableau [46] of the Runge-Kutta method (4.44)

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^T \end{array} \quad (5.14)$$

5. Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods

Substituting s by $\mathbf{s}(z)$ in $\mathbf{Z}(s) \in \mathbb{C}^{N_s \times N_s}$ results in the matrix $\mathbf{Z}(\mathbf{s}(z)) \in \mathbb{C}^{pN_s \times pN_s}$ whose computation is described in (4.89). Finally, as explained in the section 4.3.4, the product $\mathbf{Z}(s)\mathbf{J}(s)$ in the Laplace domain is transformed into a convolution between the sequences of interaction matrices $(\mathbf{Z}_i)_{i \in \mathbb{N}}$ and the current coefficients vectors $(\mathbf{J}_i)_{i \in \mathbb{N}}$ in the discrete time domain. The sequence of interaction matrices is obtained by computing the inverse Z-transform of $\mathbf{Z}(\mathbf{s}(z))$. The current and excitation vectors are evaluated for each time step i and each stage k ($k \in [1, p]$) at the times $t = t_i + \mathbf{c}_k \Delta t$. In the discrete time domain the system is

$$\sum_{j=0}^i \mathbf{Z}_j \mathbf{J}_{i-j} = \mathbf{E}_i \quad (5.15)$$

where (see (4.79),(4.75a) and (4.76))

$$\mathbf{Z}_j = \mathcal{Z}^{-1}(\mathbf{Z}(\mathbf{s}))_j \quad (5.16a)$$

$$\mathbf{J}_s(\mathbf{r}, t_i + \mathbf{c}_k \Delta t) \approx \sum_{n=1}^{N_s} [\mathbf{J}_i]_{k+(n-1)p} \mathbf{f}_n(\mathbf{r}) \quad (5.16b)$$

$$[\mathbf{E}_i]_{k+(m-1)p} = -\langle \mathbf{f}_m, \mathbf{E}^{\text{inc}}(t_i + \mathbf{c}_k \Delta t) \rangle. \quad (5.16c)$$

The MOT scheme is explicitly written as

$$\mathbf{Z}_0 \mathbf{J}_i = \mathbf{E}_i - \sum_{j=1}^i \mathbf{Z}_j \mathbf{J}_{i-j}. \quad (5.17)$$

Note that in practice the differentiated PEC-EFIE should be used instead because the non-differentiated PEC-EFIE results in a sequence of matrices that do not decay, and thus it renders the MOT unpractical. In the differentiated PEC-EFIE, the convolution on the RHS can be truncated to a finite number of terms N_{conv} . So, the MOT reads

$$\mathbf{Z}_0^{\text{der}} \mathbf{J}_i = \mathbf{E}_i^{\text{der}} - \sum_{j=1}^{N_{\text{conv}}} \mathbf{Z}_j^{\text{der}} \mathbf{J}_{i-j} \quad (5.18)$$

where the sequences of interaction matrices and excitation vectors for the differentiated PEC-EFIE are defined by

$$\mathbf{Z}_j^{\text{der}} = \mathcal{Z}^{-1}(\mathbf{Z}^{\text{der}}(\mathbf{s}))_j = \mathcal{Z}^{-1}(\tilde{\mathbf{S}}\mathbf{Z}(\mathbf{s}))_j \quad (5.19a)$$

$$[\mathbf{E}_i^{\text{der}}]_{k+(m-1)p} = -\left\langle \mathbf{f}_m, \frac{\partial \mathbf{E}^{\text{inc}}}{\partial t}(t_i + \mathbf{c}_k \Delta t) \right\rangle \quad (5.19b)$$

with

$$\mathbf{Z}^{\text{der}}(s) = s\mathbf{Z}(s) = \eta \left(-\frac{s^2}{c} \mathbf{T}_s(s) + c \mathbf{T}_h(s) \right) \quad (5.20a)$$

$$\tilde{\mathbf{S}} = \mathbf{I} \otimes \mathbf{s}. \quad (5.20b)$$

5.3 Analysis of the large time step breakdown and the DC instability

Both the PEC-EFIE and the differentiated PEC-EFIE suffer from the large time step breakdown and the DC instability. In fact, these two problems are linked to the time discretization of the equation but they are also present in other time discretization schemes such as the point testing or the space-time Galerkin. In this section, these two problems are studied for the IRK convolution quadrature.

5.3.1 Large time step breakdown

The large time step breakdown is in fact the time domain counterpart of the low frequency breakdown. In the frequency domain, it occurs because the part of the spectrum associated to the solenoidal functions scales proportionally to the frequency whereas the other part of the spectrum, associated to the non-solenoidal functions, scales proportionally to the inverse of the frequency. Overall, the relative ratio between the maximal and minimal singular values of the discrete operator is inversely proportional to the square of the frequency and so is the condition number. The scalings of the different parts of the spectrum of the discrete operator can be obtained with an Helmholtz decomposition. This decomposition is achieved using the quasi-Helmholtz projectors $\mathbf{P}^{\Lambda H}$ and $\mathbf{P}^{\Sigma} \in \mathbb{R}^{N_s \times N_s}$. These projectors are introduced in the section 1.5.3.3.

In the Laplace domain, the MoM matrix $\mathbf{Z}(s)$ is rewritten in an Helmholtz decomposed form. This is achieved by inserting the identity matrix $\mathbf{I} = \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma}$ on the two sides of $\mathbf{Z}(s)$ and rewriting the product as a block matrix form that exposes the scalings

$$\mathbf{Z}(s) = \eta (\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma}) \left(-\frac{s}{c} \mathbf{T}_s(s) + \frac{c}{s} \mathbf{T}_h(s) \right) (\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma}) \quad (5.21a)$$

$$= \eta \begin{pmatrix} \mathbf{P}^{\Lambda H} & \mathbf{P}^{\Sigma} \end{pmatrix} \begin{pmatrix} -\frac{s}{c} \mathbf{T}_s(s) & -\frac{s}{c} \mathbf{T}_s(s) \\ -\frac{s}{c} \mathbf{T}_s(s) & -\frac{s}{c} \mathbf{T}_s(s) + \frac{c}{s} \mathbf{T}_h(s) \end{pmatrix} \begin{pmatrix} \mathbf{P}^{\Lambda H} \\ \mathbf{P}^{\Sigma} \end{pmatrix} \quad (5.21b)$$

$$= \begin{pmatrix} \mathbf{P}^{\Lambda H} & \mathbf{P}^{\Sigma} \end{pmatrix} \begin{pmatrix} \mathbf{O}(s) & \mathbf{O}(s) \\ \mathbf{O}(s) & \mathbf{O}(s^{-1}) \end{pmatrix} \begin{pmatrix} \mathbf{P}^{\Lambda H} \\ \mathbf{P}^{\Sigma} \end{pmatrix} \quad (5.21c)$$

where $\mathbf{O}(x)$ denotes a matrix whose norm scales proportionally to x when $s \rightarrow 0$. In the low frequency regime, $\mathbf{T}_s(s)$ and $\mathbf{T}_h(s)$ have a static limit when $s \rightarrow 0$. Therefore, the block associated to the non-solenoidal (\mathbf{P}^{Σ}) testings and sources scales proportionally to $1/s$ while the other blocks scale proportionally to s .

Then, following the IRK convolution quadrature procedure to discretize in time, s is substituted by $\mathbf{s}(z)$ that is proportional to Δt^{-1} (5.13), and similarly $1/s$ scales proportionally to Δt after the substitution. So, the scaling of $\mathbf{Z}(\mathbf{s}(z))$ is

$$\mathbf{Z}(\mathbf{s}(z)) = \begin{pmatrix} \tilde{\mathbf{P}}^{\Lambda H} & \tilde{\mathbf{P}}^{\Sigma} \end{pmatrix} \begin{pmatrix} \mathbf{O}(\Delta t^{-1}) & \mathbf{O}(\Delta t^{-1}) \\ \mathbf{O}(\Delta t^{-1}) & \mathbf{O}(\Delta t) \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{P}}^{\Lambda H} \\ \tilde{\mathbf{P}}^{\Sigma} \end{pmatrix}. \quad (5.22)$$

5. Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods

where $\tilde{\mathbf{P}}^{\Lambda H} = \mathbf{P}^{\Lambda H} \otimes \mathbf{I}_p$ and $\tilde{\mathbf{P}}^{\Sigma} = \mathbf{P}^{\Sigma} \otimes \mathbf{I}_p$ are the projectors that have been modified to be made compatible with the matrices in $\mathbb{C}^{pN_s \times pN_s}$ (see section 4.3.5.3). Finally, the inverse Z-transform does not affect the scalings, so \mathbf{Z}_0 also scales as

$$\mathbf{Z}_0 = \begin{pmatrix} \tilde{\mathbf{P}}^{\Lambda H} & \tilde{\mathbf{P}}^{\Sigma} \end{pmatrix} \begin{pmatrix} \mathbf{O}(\Delta t^{-1}) & \mathbf{O}(\Delta t^{-1}) \\ \mathbf{O}(\Delta t^{-1}) & \mathbf{O}(\Delta t) \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{P}}^{\Lambda H} \\ \tilde{\mathbf{P}}^{\Sigma} \end{pmatrix}. \quad (5.23)$$

Therefore, the condition number of \mathbf{Z}_0 scales proportionally to Δt^2 when $\Delta t \rightarrow +\infty$.

5.3.2 DC instability

5.3.2.1 DC instability in the PEC-EFIE

As explained in the section 4.4, the DC instability is caused by a constant solenoidal current that creeps into the solution. By definition a constant solenoidal current \mathbf{J}_s^{cs} verifies $\frac{\partial \mathbf{J}_s^{\text{cs}}}{\partial t} = \mathbf{0}$ and $\nabla \cdot \mathbf{J}_s^{\text{cs}} = 0$. Therefore, it is in the nullspace of \mathcal{T} (i.e. $\mathcal{T}\mathbf{J}_s^{\text{cs}}(\mathbf{r}, t) = \mathbf{0}$ for all \mathbf{r} and t). With the initial condition $\mathbf{J}_s^{\text{cs}}(\mathbf{r}, t = 0) = \mathbf{0}$ (which is assumed implicitly by imposing that the solution of the PEC-EFIE verifies $\mathbf{J}_s(\mathbf{r}, t) = \mathbf{0}$ for all $t < 0$), the equation

$$\frac{\partial}{\partial t} \mathbf{J}_s^{\text{cs}}(\mathbf{r}, t) = \mathbf{0} \quad (5.24)$$

is transformed into the Laplace domain equation

$$s\mathbf{J}_s^{\text{cs}}(\mathbf{r}, s) = \mathbf{0}. \quad (5.25)$$

Following the IRK convolution quadrature procedure for the time discretization, the factor s is substituted by $\tilde{\mathbf{s}} = \mathbf{I} \otimes \mathbf{s}$ (see section 4.3.5.3). Also, the spurious current is discretized similarly to (5.16b) by the sequence $(\mathbf{J}_i^{\text{cs}})_i$ such that

$$\mathbf{J}_s^{\text{cs}}(\mathbf{r}, t_i + \mathbf{c}_k \Delta t) \approx \sum_{n=1}^{N_s} [\mathbf{J}_i^{\text{cs}}]_{k+(n-1)p} \mathbf{f}_n(\mathbf{r}). \quad (5.26)$$

Then, in the discrete time domain, the sequence $(\mathbf{J}_i^{\text{cs}})_i$ verifies for all $i > 0$

$$\sum_{j=0}^i \mathcal{Z}^{-1}(\tilde{\mathbf{s}})_j \mathbf{J}_{i-j}^{\text{cs}} = \mathbf{0}. \quad (5.27)$$

Initializing the sequence with $\mathbf{J}_0^{\text{cs}} = \mathbf{0}$ should result in $\mathbf{J}_i^{\text{cs}} = \mathbf{0}$ for all i . But in practice, there are numerical errors that produce a non-zero constant solenoidal part in the solution. Using the L-stability [47] of the IRK method, $\mathbf{s}(z)$ is rewritten as in (4.56) to involve a finite number of power of z^{-1} .

$$\mathbf{s}(z) = \frac{1}{\Delta t} (\mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{1} \mathbf{b}^{\top} \mathbf{A}^{-1} z^{-1}). \quad (5.28)$$

As it can be read in (4.39a), z^{-1} in the Z-domain actually corresponds to taking the previous element in the time domain sequence. Therefore, the inverse Z-transform of $\mathbf{s}(z)$ is defined by

$$\mathcal{Z}^{-1}(\mathbf{s})_i = \frac{1}{\Delta t} (\mathbf{A}^{-1}\delta_{i,0} - \mathbf{A}^{-1}\mathbf{1b}^T\mathbf{A}^{-1}\delta_{i-1,0}) \quad (5.29)$$

where δ denotes the Kronecker delta. This expression is inserted in (5.27) which results in the following recurrence relation

$$\mathbf{J}_i^{\text{cs}} = \widetilde{\mathbf{1b}}^T \widetilde{\mathbf{A}}^{-1} \mathbf{J}_{i-1}^{\text{cs}} \quad (5.30)$$

where $\widetilde{\mathbf{A}} = \mathbf{I} \otimes \mathbf{A}$ and $\widetilde{\mathbf{1b}}^T = \mathbf{I} \otimes (\mathbf{1b}^T)$. But in practice, the non-differentiated PEC-EFIE is not used. Instead the time differentiated PEC-EFIE is solved, and it has a different characterization of the DC current.

5.3.2.2 DC instability in the differentiated PEC-EFIE

In the differentiated PEC-EFIE, the DC instability is worse because in addition to the constant solenoidal currents, also the linear in time solenoidal currents are in the nullspace of the operator $\frac{\partial}{\partial t}\mathcal{T}$. Let \mathbf{J}_s^{ls} be such a linear in time solenoidal current.

It verifies $\frac{\partial^2 \mathbf{J}_s^{\text{ls}}}{\partial t^2} = \mathbf{0}$ and $\nabla \cdot \mathbf{J}_s^{\text{ls}} = 0$. Therefore, $\frac{\partial}{\partial t}\mathcal{T}\mathbf{J}_s^{\text{ls}} = \mathbf{0}$. With the initial conditions $\mathbf{J}_s^{\text{ls}}(\mathbf{r}, t=0) = \mathbf{0}$ and $\frac{\partial}{\partial t}\mathbf{J}_s^{\text{ls}}(\mathbf{r}, t=0) = \mathbf{0}$, the equation

$$\frac{\partial^2}{\partial t^2} \mathbf{J}_s^{\text{ls}}(\mathbf{r}, t) = \mathbf{0} \quad (5.31)$$

is transformed into the Laplace domain equation

$$s^2 \mathbf{J}_s^{\text{ls}}(\mathbf{r}, s) = \mathbf{0}. \quad (5.32)$$

Similarly to the previous section, after the time discretization, the time domain sequence of the spurious currents $(\mathbf{J}_i^{\text{ls}})_i$ verifies for all $i > 1$

$$\sum_{j=0}^i \mathcal{Z}^{-1}(\widetilde{\mathbf{s}}^2)_j \mathbf{J}_{i-j}^{\text{ls}} = \mathbf{0}. \quad (5.33)$$

Again, with the simplified expression of $\mathbf{s}(z)$ (5.28), the recurrence equation (5.33) becomes

$$\mathbf{J}_i^{\text{ls}} = \widetilde{\mathbf{A}} \left(\widetilde{\mathbf{A}}^{-1} \widetilde{\mathbf{1b}}^T + \widetilde{\mathbf{1b}}^T \widetilde{\mathbf{A}}^{-1} \right) \widetilde{\mathbf{A}}^{-1} \mathbf{J}_{i-1}^{\text{ls}} - \widetilde{\mathbf{A}}^2 \left(\widetilde{\mathbf{A}}^{-1} \widetilde{\mathbf{1b}}^T \widetilde{\mathbf{A}}^{-1} \right)^2 \mathbf{J}_{i-2}^{\text{ls}} \quad (5.34)$$

Note that these characterizations of the DC currents (5.30) and (5.34) do not depend on the time step Δt .

5.4 Regularized PEC-EFIE

In this section the regularized PEC-EFIE (without DC instability and stable for large time steps) is constructed using the quasi-Helmholtz projectors.

5.4.1 Regularization in the Laplace domain

The DC instability is due to the presence of constant (or linear in time for the differentiated PEC-EFIE) solenoidal currents in the nullspace of the operator. Therefore, by integrating in time the solenoidal part (ΛH) of the operator, the time derivative is effectively removed and the only constant current allowed in the nullspace is $\mathbf{0}$. Then, to remove the time integration in the non-solenoidal (Σ) part of the operator, the non-solenoidal part is differentiated. In the Laplace domain, a time derivative corresponds to a multiplication by s and an integration in time corresponds to a multiplication by $1/s$. Therefore, the system (5.7) is regularized in the Laplace domain as

$$\mathbf{L}(s)\mathbf{Z}(s)\mathbf{R}(s)\mathbf{Y}(s) = \mathbf{L}(s)\mathbf{E}(s) \quad (5.35)$$

where $\mathbf{Y}(s)$ is an auxiliary unknown such that

$$\mathbf{J}(s) = \mathbf{R}(s)\mathbf{Y}(s) \quad (5.36)$$

and where $\mathbf{L}(s)$ and $\mathbf{R}(s)$ are the preconditioners

$$\mathbf{L}(s) = \frac{c}{sa}\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \quad (5.37a)$$

$$\mathbf{R}(s) = \mathbf{P}^{\Lambda H} + \frac{sa}{c}\mathbf{P}^{\Sigma}. \quad (5.37b)$$

Here, the speed of light c was also used in the preconditioner to further lower the ill-scaling between the two parts of the spectrum. In addition, a is a length parameter that ensures a consistent dimensionality. a should be chosen proportional to the diameter of Γ so that a spacial rescaling of the simulation (i.e. that rescales equally the geometry and the wavelength) does not modify the conditioning of the system. Note that the definitions of the Laplace domain $\mathbf{L}(s)$ and $\mathbf{R}(s)$ in (5.37) are consistent with the preconditioners of the Fourier domain PEC-EFIE in (1.196). The asymptotic behaviour of the regularized system noted $\mathbf{Z}^{\text{reg}}(s)$ can be visualized in a block matrix

form similar to (5.21b)

$$\mathbf{Z}^{\text{reg}}(s) = \mathbf{L}(s)\mathbf{Z}(s)\mathbf{R}(s) \quad (5.38a)$$

$$\begin{aligned} &= \eta \left(-\frac{1}{a}\mathbf{P}^{\Lambda H}\mathbf{T}_s(s)\mathbf{P}^{\Lambda H} + a\mathbf{T}_h(s) - \frac{s^2a}{c^2}\mathbf{P}^{\Sigma}\mathbf{T}_s(s)\mathbf{P}^{\Sigma} \right. \\ &\quad \left. - \frac{s}{c}\mathbf{P}^{\Lambda H}\mathbf{T}_s(s)\mathbf{P}^{\Sigma} - \frac{s}{c}\mathbf{P}^{\Sigma}\mathbf{T}_s(s)\mathbf{P}^{\Lambda H} \right) \end{aligned} \quad (5.38b)$$

$$= \eta \begin{pmatrix} \mathbf{P}^{\Lambda H} & \mathbf{P}^{\Sigma} \end{pmatrix} \begin{pmatrix} -\frac{1}{a}\mathbf{T}_s(s) & -\frac{s}{c}\mathbf{T}_s(s) \\ -\frac{s}{c}\mathbf{T}_s(s) & -\frac{s^2a}{c^2}\mathbf{T}_s(s) + a\mathbf{T}_h(s) \end{pmatrix} \begin{pmatrix} \mathbf{P}^{\Lambda H} \\ \mathbf{P}^{\Sigma} \end{pmatrix} \quad (5.38c)$$

$$= \eta \begin{pmatrix} \mathbf{P}^{\Lambda H} & \mathbf{P}^{\Sigma} \end{pmatrix} \begin{pmatrix} -\frac{1}{a}\mathbf{T}_s(s) & \mathbf{0} \\ \mathbf{0} & a\mathbf{T}_h(s) \end{pmatrix} \begin{pmatrix} \mathbf{P}^{\Lambda H} \\ \mathbf{P}^{\Sigma} \end{pmatrix} + \mathbf{O}_{s \rightarrow 0}(s) \quad (5.38d)$$

where $\mathbf{O}_{s \rightarrow 0}(s)$ is a matrix whose norm scales proportionally to s when $s \rightarrow 0$. So, $\mathbf{Z}^{\text{reg}}(s)$ has a limit when $s \rightarrow 0$ which is $\mathbf{Z}^{\text{reg}}(0) = -\eta\mathbf{P}^{\Lambda H}\mathbf{T}_s(0)\mathbf{P}^{\Lambda H}/a + a\mathbf{T}_h(0)$.

The regularized system in the Laplace domain is rewritten as

$$\mathbf{Z}^{\text{reg}}(s)\mathbf{Y}(s) = \mathbf{V}(s) \quad (5.39)$$

where the RHS is

$$\mathbf{V}(s) = \left(\frac{c}{sa}\mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \right) \mathbf{E}(s) \quad (5.40)$$

and the original unknown is retrieved as

$$\mathbf{J}(s) = \left(\mathbf{P}^{\Lambda H} + \frac{sa}{c}\mathbf{P}^{\Sigma} \right) \mathbf{Y}(s). \quad (5.41)$$

5.4.2 Time discretization of the regularized system

5.4.2.1 Interaction matrices

Following the procedure for the IRK convolution quadrature discretization, the scalar s is substituted by the matrix $\mathbf{s}(z)$ in $\mathbf{Z}^{\text{reg}}(s)$. The matrix \mathbf{s} is diagonalized as $\mathbf{s} = \mathbf{MDM}^{-1}$. The $N_s \times N_s$ identity is noted \mathbf{I} and the $p \times p$ identity is noted \mathbf{I}_p . Then, using the expanded expression of $\mathbf{Z}^{\text{reg}}(s)$ in (5.38b), $\mathbf{Z}^{\text{reg}}(\mathbf{s}(z))$ is

$$\mathbf{Z}^{\text{reg}}(\mathbf{s}) = \eta \left(-\frac{1}{a}\tilde{\mathbf{P}}^{\Lambda H}\tilde{\mathbf{M}}\mathbf{T}_s(\mathbf{D})\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{P}}^{\Lambda H} \right. \quad (5.42a)$$

$$\left. + a\tilde{\mathbf{M}}\mathbf{T}_h(\mathbf{D})\tilde{\mathbf{M}}^{-1} \right. \quad (5.42b)$$

$$\left. - \frac{a}{c^2}\tilde{\mathbf{P}}^{\Sigma}\tilde{\mathbf{M}}\tilde{\mathbf{D}}^2\mathbf{T}_s(\mathbf{D})\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{P}}^{\Sigma} \right. \quad (5.42c)$$

$$\left. - \frac{1}{c}\tilde{\mathbf{P}}^{\Lambda H}\tilde{\mathbf{M}}\tilde{\mathbf{D}}\mathbf{T}_s(\mathbf{D})\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{P}}^{\Sigma} \right. \quad (5.42d)$$

$$\left. - \frac{1}{c}\tilde{\mathbf{P}}^{\Sigma}\tilde{\mathbf{M}}\tilde{\mathbf{D}}\mathbf{T}_s(\mathbf{D})\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{P}}^{\Lambda H} \right) \quad (5.42e)$$

5. Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods

where $\tilde{\mathbf{s}} = \mathbf{I} \otimes \mathbf{s}$, $\tilde{\mathbf{M}} = \mathbf{I} \otimes \mathbf{M}$, $\tilde{\mathbf{D}} = \mathbf{I} \otimes \mathbf{D}$, $\tilde{\mathbf{P}}^{\Lambda H} = \mathbf{P}^{\Lambda H} \otimes \mathbf{I}_p$ and $\tilde{\mathbf{P}}^{\Sigma} = \mathbf{P}^{\Sigma} \otimes \mathbf{I}_p$. Details on this transformation can be found in the section 4.3.5.3. This expression of $\mathbf{Z}^{\text{reg}}(\mathbf{s})$ (5.42) represents how it is computed in practice. Finally, the sequence of interaction matrices $(\mathbf{Z}_j^{\text{reg}})_j$ is obtained with an inverse Z-transform (4.37) of $\mathbf{Z}^{\text{reg}}(\mathbf{s})$

$$\mathbf{Z}_j^{\text{reg}} = \mathcal{Z}^{-1}(\mathbf{Z}^{\text{reg}}(\mathbf{s}))_j. \quad (5.43)$$

Similarly to (4.81), the norm of the stabilized interaction matrices decreases exponentially i.e.

$$\|\mathbf{Z}_j^{\text{reg}}\| \xrightarrow{j \rightarrow +\infty} 0 \text{ exponentially.} \quad (5.44)$$

5.4.2.2 Excitation vector (analytic evaluation)

The solenoidal part of the RHS $\mathbf{V}(s)$ in (5.40) contains a $1/s$ factor. $1/s$ in the Laplace domain corresponds to an integration in the time domain. Therefore, a primitive of \mathbf{E}^{inc} noted \mathbf{E}^{prim} must be computed

$$\frac{\partial \mathbf{E}^{\text{prim}}}{\partial t}(\mathbf{r}, t) = \mathbf{E}^{\text{inc}}(\mathbf{r}, t). \quad (5.45)$$

In the context of the IRK convolution quadrature, the analogue of the Laplace domain RHS $\mathbf{V}(s)$ (5.40) is a sequence $(\mathbf{V}_i)_i$ in the discrete time domain

$$\mathbf{V}_i = \frac{c}{a} \tilde{\mathbf{P}}^{\Lambda H} \mathbf{E}_i^{\text{prim}} + \tilde{\mathbf{P}}^{\Sigma} \mathbf{E}_i \quad (5.46)$$

where $(\mathbf{E}_i^{\text{prim}})_i$ is a sequence defined similarly to $(\mathbf{E}_i)_i$ in (5.16c) but using a primitive of \mathbf{E}^{inc} (noted \mathbf{E}^{prim}) as a source

$$[\mathbf{E}_i^{\text{prim}}]_{k+(m-1)p} = -\langle \mathbf{f}_m, \mathbf{E}^{\text{prim}}(t_i + \mathbf{c}_k \Delta t) \rangle. \quad (5.47)$$

Note that the choice of the primitive constant for \mathbf{E}^{prim} does not modify the value of \mathbf{V}_i because it is cancelled by the multiplication with $\tilde{\mathbf{P}}^{\Lambda H}$. Indeed, from (1.153), the testing of a constant field by solenoidal functions is 0. Though, this primitive constant must be chosen such that $\lim_{t \rightarrow \pm\infty} \mathbf{E}^{\text{prim}}(\mathbf{r}, t) = \mathbf{0}$ to compute \mathbf{V}_i accurately in the early and the late time steps. This is done to avoid numerical cancellations in \mathbf{V}_i . In fact, unless \mathbf{E}^{inc} is an odd function, there is no guarantee that $\lim_{t \rightarrow -\infty} \mathbf{E}^{\text{prim}}(\mathbf{r}, t) = \lim_{t \rightarrow +\infty} \mathbf{E}^{\text{prim}}(\mathbf{r}, t)$. In this case different primitives are used for the computation of the sequence $(\mathbf{E}_i^{\text{prim}})_i$ in the early time steps and the late time steps.

5.4.2.3 Excitation vector (numerical evaluation)

If no analytic expression for a primitive of $\mathbf{E}^{\text{inc}}(\mathbf{r}, t)$ is available, then the sequence of excitation vectors $(\mathbf{E}_i^{\text{prim}})_i$ should be computed numerically from the sequence $(\mathbf{E}_i)_i$ using the IRK scheme. The Laplace domain equation corresponding to (5.45) is

$$s \mathbf{E}^{\text{prim}}(\mathbf{r}, s) = \mathbf{E}^{\text{inc}}(\mathbf{r}, s). \quad (5.48)$$

After the discretization in space and the substitution of s by $\mathbf{s}(z)$, the following equation in the Z -domain holds

$$\tilde{\mathbf{s}}(z)\mathcal{Z}(\mathbf{E}^{\text{prim}})(z) = \mathcal{Z}(\mathbf{E})(z). \quad (5.49)$$

Expanding $\mathbf{s}(z)$ (5.13) and rearranging the terms yields

$$(z-1)\mathcal{Z}(\mathbf{E}^{\text{prim}})(z) = \Delta t \left((z-1)\tilde{\mathbf{A}} + \tilde{\mathbf{1b}}^\top \right) \mathcal{Z}(\mathbf{E})(z). \quad (5.50)$$

Using (4.39) results in

$$\mathbf{E}_{i+1}^{\text{prim}} + \delta_{i+1,0}\mathbf{E}_0^{\text{prim}} - \mathbf{E}_i^{\text{prim}} = \Delta t \left(\tilde{\mathbf{A}}(\mathbf{E}_{i+1} + \delta_{i+1,0}\mathbf{E}_0 - \mathbf{E}_i) + \tilde{\mathbf{1b}}^\top \mathbf{E}_i \right) \quad (5.51)$$

which gives the initial condition for $i = -1$ and the recurrence relation for all $i \in \mathbb{N}$

$$\mathbf{E}_0^{\text{prim}} = \Delta t \tilde{\mathbf{A}} \mathbf{E}_0 \quad (5.52a)$$

$$\mathbf{E}_{i+1}^{\text{prim}} = \mathbf{E}_i^{\text{prim}} + \Delta t \left(\tilde{\mathbf{A}}(\mathbf{E}_{i+1} - \mathbf{E}_i) + \tilde{\mathbf{1b}}^\top \mathbf{E}_i \right). \quad (5.52b)$$

This is the forward computation of the excitation vector. It is used in the early time steps, but similarly to the analytic computation, there is not guarantee that the primitive goes to $\mathbf{0}$ for $t \rightarrow +\infty$. Actually, due to the iterative nature of the computation, there would still be a saturation in the order of the machine precision due to numerical errors even if the primitive goes to $\mathbf{0}$.

To get a primitive that cancels in late time steps, a backward computation also has to be performed. Assuming that the simulation has to be computed for N_t time steps, the backward computation is given by the following recurrence relation for all $i \leq N_t$

$$\mathbf{E}_{N_t}^{\text{prim}} = \Delta t \tilde{\mathbf{A}} \mathbf{E}_{N_t} \quad (5.53a)$$

$$\mathbf{E}_{i-1}^{\text{prim}} = \mathbf{E}_i^{\text{prim}} + \Delta t \left(\tilde{\mathbf{A}}(\mathbf{E}_{i-1} - \mathbf{E}_i) - \tilde{\mathbf{1b}}^\top \mathbf{E}_{i-1} \right). \quad (5.53b)$$

5.4.2.4 Solution

Applying the IRK convolution quadrature procedure to (5.39) results in the following convolution

$$\sum_{j=0}^i \mathbf{z}_j^{\text{reg}} \mathbf{Y}_{i-j} = \mathbf{V}_i \quad (5.54)$$

where $(\mathbf{Y}_i)_i$ is the sequence of vectors in \mathbb{R}^{pN_s} that corresponds to the auxiliary unknown whose relation to the original unknown $(\mathbf{J}_i)_i$ is given in the Laplace domain in (5.41). The convolution is rewritten to make apparent the regularized MOT scheme

$$\mathbf{z}_0^{\text{reg}} \mathbf{Y}_i = \mathbf{V}_i - \sum_{j=1}^{N_{\text{conv}}} \mathbf{z}_j^{\text{reg}} \mathbf{Y}_{i-j} \quad (5.55)$$

5. Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods

where N_{conv} has been introduced to truncate the convolution to a finite number of terms at each time step because the norm of the interaction matrices $\mathbf{Z}_j^{\text{reg}}$ decreases exponentially fast with increasing indices j (5.44).

After the completion of the MOT, the auxiliary current $(\mathbf{Y}_i)_{i=0}^{N_t}$ is known and it must be used to reconstruct the original unknown $(\mathbf{J}_i)_{i=0}^{N_t}$. Substituting s by $\mathbf{s}(z)$ in (5.41) and inverse Z-transforming it results in the following relation for all $i \in [0, N_t]$

$$\mathbf{J}_i = \tilde{\mathbf{P}}^{\Lambda H} \mathbf{Y}_i + \frac{a}{c} \tilde{\mathbf{P}}^{\Sigma} \sum_{j=0}^i \mathcal{Z}^{-1}(\tilde{\mathbf{s}}) \mathbf{Y}_{i-j}. \quad (5.56)$$

Using the expression of $\mathcal{Z}^{-1}(\mathbf{s})$ in (5.29), the above equation simplifies to

$$\mathbf{J}_i = \tilde{\mathbf{P}}^{\Lambda H} \mathbf{Y}_i + \frac{a}{c \Delta t} \tilde{\mathbf{P}}^{\Sigma} \tilde{\mathbf{A}}^{-1} \left(\mathbf{Y}_i - \tilde{\mathbf{1}} \mathbf{b}^T \tilde{\mathbf{A}}^{-1} \mathbf{Y}_{i-1} \right). \quad (5.57)$$

This sequence of vectors is used to evaluate the current $\mathbf{J}_s(\mathbf{r}, t)$ using (5.16b).

5.5 Implementation details

This section describes how the formulation presented in this chapter can be solved in practice. In particular, it is explained why this formulation is compatible with the existing fast techniques to solve integral equations.

5.5.1 Quasi-Helmholtz projectors

As it is mentioned in the section 1.5.3.3, the quasi-Helmholtz projectors $\mathbf{P}^{\Lambda H}$ and \mathbf{P}^{Σ} can be multiplied by a vector in a linear complexity $\mathcal{O}(N_s)$ [38]. But in this formulation the projectors are used in a Kronecker product with the $p \times p$ identity matrix as $\tilde{\mathbf{P}}^{\Lambda H} = \mathbf{P}^{\Lambda H} \otimes \mathbf{I}_p$ and $\tilde{\mathbf{P}}^{\Sigma} = \mathbf{P}^{\Sigma} \otimes \mathbf{I}_p$.

Consider a vector $\mathbf{X} \in \mathbb{C}^{pN_s}$. The elements of this vector are stored in p subvectors \mathbf{x}^k in \mathbb{C}^{N_s} such that for all $k \in [1, p]$ and $m \in [1, N_s]$,

$$[\mathbf{x}^k]_m = \mathbf{X}_{p(m-1)+k}. \quad (5.58)$$

Then, the product of an augmented projector by \mathbf{X} is computed as

$$[\tilde{\mathbf{P}}^{\Sigma} \mathbf{X}]_{p(m-1)+k} = [\mathbf{P}^{\Sigma} \mathbf{x}^k]_m. \quad (5.59)$$

In other words, the multiplication of a vector by $\tilde{\mathbf{P}}^{\Sigma}$ can be computed with p multiplications of \mathbf{P}^{Σ} i.e. the cost is $\mathcal{O}(N_s)$.

5.5.2 Excitation vectors

The vectors \mathbf{E}_i and $\mathbf{E}_i^{\text{prim}}$ are computed for each time step $i \in [0, N_t]$. With the analytic evaluation, each vector contains pN_s elements, and each element is computed in a constant time, so the total cost is $\mathcal{O}(N_s N_t)$.

Instead, if the vectors $\mathbf{E}_i^{\text{prim}}$ are computed sequentially with the numerical evaluation as described in the section 5.4.2.3, then they require the products of the matrices $\tilde{\mathbf{A}}$ or $\tilde{\mathbf{1b}}^\top$ by a vector. Since these matrices are block diagonal, they are sparse and their products by a vector are computed in a linear complexity, so the complexity is still $\mathcal{O}(N_s N_t)$.

Note that in this case of the numerical evaluation of the excitation vectors, the fact the computation of the vectors in the late time steps is performed backward requires their storage which costs $\mathcal{O}(N_s N_t)$ in memory, as opposed to the analytic evaluation where the vectors can be computed independently.

5.5.3 Retrieving the original current from the auxiliary current

Once the sequence of auxiliary unknowns $(\mathbf{Y}_i)_i$ are found after the completion of the MOT, the original current sequence $(\mathbf{J}_i)_i$ is computed using (5.57). At each time step $i \in [0, N_t]$, two multiplications of vectors with the augmented projectors $\tilde{\mathbf{P}}^{\Lambda H}$ and $\tilde{\mathbf{P}}^\Sigma$ are needed. These multiplications are $\mathcal{O}(N_s)$ in complexity as explain above. Three multiplications with the block diagonal matrices $\tilde{\mathbf{A}}^{-1}$ and $\tilde{\mathbf{1b}}^\top$ are needed and they are also linear in complexity. So, the total cost of retrieving the original current is $\mathcal{O}(N_s N_t)$.

5.5.4 Marching-on-in-time

At each time step $i \in [0, N_t]$, the linear system in (5.55) is solved for \mathbf{Y}_i . This system is solved iteratively and it can be assumed that it requires a fixed number of iterations N_{iter} to converge. Each iteration involves the multiplication of $\mathbf{Z}_0^{\text{reg}}$ by a vector. Also, the convolution requires N_{conv} multiplications of $\mathbf{Z}_j^{\text{reg}}$ by a vector. So overall, there are $\mathcal{O}((N_{\text{iter}} + N_{\text{conv}})N_t)$ multiplications of the interaction matrices $\mathbf{Z}_j^{\text{reg}}$ by a vector.

5.5.5 Interaction matrix multiplication by a vector

Inserting the product of $\mathbf{Z}_j^{\text{reg}}$ by a vector \mathbf{X} in (5.43) and using the trapezoidal rules (4.38b) to compute the inverse Z-transform yields

$$\mathbf{Z}_j^{\text{reg}}\mathbf{X} = \frac{\rho^n}{Q} \sum_{q=0}^{Q-1} e^{2\pi i \frac{q}{Q} j} \mathbf{Z}^{\text{reg}} \left(\mathbf{s} \left(\rho e^{2\pi i \frac{q}{Q}} \right) \right) \mathbf{X}. \quad (5.60)$$

So the computation of $\mathbf{Z}_j^{\text{reg}}\mathbf{X}$ requires Q multiplications of $\mathbf{Z}^{\text{reg}}(\mathbf{s})$ by \mathbf{X} . In fact, since the temporal sequence is real, half of the multiplications in the inverse-Z transform are actually required by taking advantage of the complex conjugation of the terms in the sum

$$\begin{aligned} 2\Re \left(e^{2\pi i \frac{q}{Q} j} \mathbf{Z}^{\text{reg}} \left(\mathbf{s} \left(\rho e^{2\pi i \frac{q}{Q}} \right) \right) \mathbf{X} \right) &= e^{2\pi i \frac{q}{Q} j} \mathbf{Z}^{\text{reg}} \left(\mathbf{s} \left(\rho e^{2\pi i \frac{q}{Q}} \right) \right) \mathbf{X} \\ &+ e^{2\pi i \frac{Q-q}{Q} j} \mathbf{Z}^{\text{reg}} \left(\mathbf{s} \left(\rho e^{2\pi i \frac{Q-q}{Q}} \right) \right) \mathbf{X}. \end{aligned} \quad (5.61)$$

5. Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods

It can be read from (5.42) that the computation of $\mathbf{Z}^{\text{reg}}(\mathbf{s})\mathbf{X}$ involves multiplications with the augmented projectors $\tilde{\mathbf{P}}^\Sigma$ and $\tilde{\mathbf{P}}^{\Lambda H}$ (these can actually be factorized out of the Z-transform since they do not depend on z), and with the block diagonal matrices $\tilde{\mathbf{M}}$, $\tilde{\mathbf{M}}^{-1}$ and $\tilde{\mathbf{D}}$. These multiplications all have a linear complexity $\mathcal{O}(N_s)$ as explained before. Finally, there are also five multiplications of $\mathbf{T}_s(\mathbf{D})$ or $\mathbf{T}_h(\mathbf{D})$ by a vector. Similarly to the multiplication with the projectors, a product in the form $\mathbf{T}_s(\mathbf{D})\mathbf{X}$ can be computed from p products in the form $\mathbf{T}_s(\mathbf{D}_{kk})\mathbf{x}^k$ where $\mathbf{T}_s(\mathbf{D}_{kk}) \in \mathbb{C}^{pN_s \times pN_s}$ is the interaction matrix in the Laplace domain evaluated at the frequency \mathbf{D}_{kk} , and $\mathbf{x}^k \in \mathbb{C}^{N_s}$ ($k \in [1, p]$) are p vectors built from the elements of $\mathbf{X} \in \mathbb{C}^{pN_s}$. From (5.58) and (4.90), it results for all $k \in [1, p]$ and $m \in [1, N_s]$

$$[\mathbf{T}_s(\mathbf{D})\mathbf{X}]_{p(m-1)+k} = [\mathbf{T}_s(\mathbf{D}_{kk})\mathbf{x}^k]_m. \quad (5.62)$$

In other words, $\mathbf{T}_s(\mathbf{D})\mathbf{X}$ can be computed with p multiplications of $\mathbf{T}_s(s)$ by a vector. The Multi-Level Fast Multipole Method (MLFMM) [26, 27] is a fast algorithm for these multiplications that has a complexity $\mathcal{O}(N_s \log(N_s))$. The same result holds for \mathbf{T}_h .

To summarize, there are $\mathcal{O}((N_{\text{iter}} + N_{\text{conv}})N_t)$ computations in the form $\mathbf{Z}_j^{\text{reg}}\mathbf{X}$ and each costs $\mathcal{O}(QN_s \log(N_s))$. So overall, the complexity of the solver is $\mathcal{O}(Q(N_{\text{iter}} + N_{\text{conv}})N_t N_s \log(N_s))$ which is the dominant complexity.

Note that if the system is well-conditioned, the number of iterations N_{iter} is likely to be a low constant. Also, the number of interaction matrices N_{conv} required in the convolution on the RHS of the MOT, as well as the number of terms Q required in the computation of the inverse Z-transform in (4.38b) decrease with the frequency.

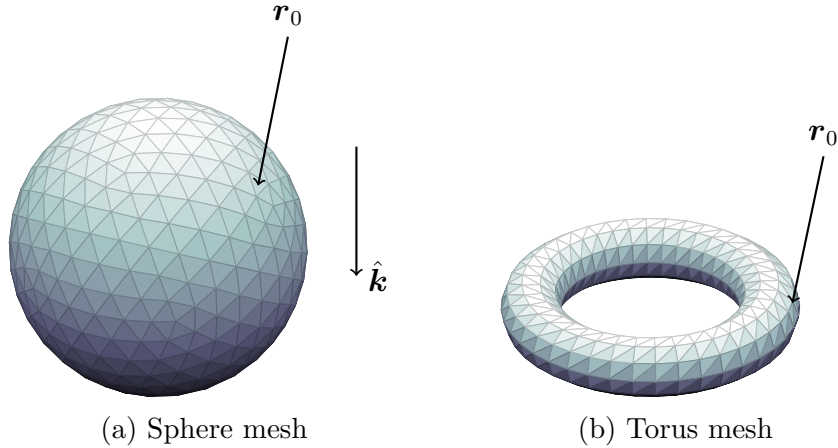
5.6 Numerical results

In these numerical results, the excitation field $\mathbf{E}^{\text{inc}}(\mathbf{r}, t)$ is a modulated Gaussian plane wave whose expression is

$$\mathbf{E}^{\text{inc}}(\mathbf{r}, t) = \hat{\mathbf{p}} \exp\left(-\frac{\tau_d(\mathbf{r}, t)^2}{2\tau_{\text{ch}}^2}\right) \cos(2\pi f_0 \tau_d(\mathbf{r}, t)) E_0 \quad (5.63)$$

where

- $\hat{\mathbf{k}}$ is the direction of propagation,
- $\tau_d(\mathbf{r}, t) = t - \frac{1}{c} \hat{\mathbf{k}} \cdot \mathbf{r}$ is the delayed time (in s),
- $\hat{\mathbf{p}}$ is the polarization,
- E_0 is the peak amplitude (in V/m),
- f_0 is the central frequency (in Hz),
- τ_{ch} is the characteristic time (in s) that depends on the frequency bandwidth f_{bw} as $\tau_{\text{ch}} = \frac{6}{2\pi f_{\text{bw}}}$.

Figure 5.1: Meshes and positions of the current probe \mathbf{r}_0

As explained in the section 5.2.1, this incident field is assumed to be $\mathbf{0}$ in the neighbourhood of Γ for all $t < t_0$. Note that this is equivalent to assume that the field is $\mathbf{0}$ for all $t < 0$ with $\tau_d(\mathbf{r}, t) = t - t_0 - \frac{1}{c} \hat{\mathbf{k}} \cdot \mathbf{r}$.

For the computation of the RHS in the regularized formulation, two primitives of \mathbf{E}^{inc} are needed (see section 5.4.2.2). They are given by the following expressions

$$\mathbf{E}_-^{\text{prim}}(\mathbf{r}, t) = \hat{\mathbf{p}} \alpha \Re(\text{erfc}(-\beta(\mathbf{r}, t))) \quad (5.64a)$$

$$\mathbf{E}_+^{\text{prim}}(\mathbf{r}, t) = -\hat{\mathbf{p}} \alpha \Re(\text{erfc}(\beta(\mathbf{r}, t))) \quad (5.64b)$$

$$\alpha = \sqrt{\frac{\pi}{2}} \tau_{\text{ch}} \exp(-2\pi^2 f_0^2 \tau_{\text{ch}}^2) E_0 \quad (5.64c)$$

$$\beta(\mathbf{r}, t) = \frac{\tau_d(\mathbf{r}, t) + 2\pi i f_0 \tau_{\text{ch}}^2}{\sqrt{2} \tau_{\text{ch}}} \quad (5.64d)$$

where $\Re(\text{erfc}(\cdot))$ is the real part of the complementary error function. They verify

$$\lim_{t \rightarrow -\infty} \mathbf{E}_-^{\text{prim}}(\mathbf{r}, t) = \mathbf{0} \quad (5.65a)$$

$$\lim_{t \rightarrow +\infty} \mathbf{E}_+^{\text{prim}}(\mathbf{r}, t) = \mathbf{0}. \quad (5.65b)$$

Finally, a set of parameters have been kept constant in these simulations. The excitation plane wave in (5.63) has an amplitude $E_0 = 1$ V/m, a direction of propagation $\hat{\mathbf{k}} = -\hat{\mathbf{z}}$ and a polarization $\hat{\mathbf{p}} = \hat{\mathbf{x}}$. The IRK method used is the 3 stages Radau IIA method whose Butcher tableau is in (4.45). For the computation of the inverse Z-transform, $Q = 32$ and $\rho = 1.1$ have been used in the trapezoidal rules (4.38b). The length parameter a that appears in the preconditioners (5.37) is set to 1 meter.

5.6.1 Results related to the DC instability

The figure 5.2 shows the current density on a unit sphere discretized with $N_s = 1080$ edges whose average length is $h \approx 0.20$ m. The simulation is done on $N_t = 500$

5. Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods

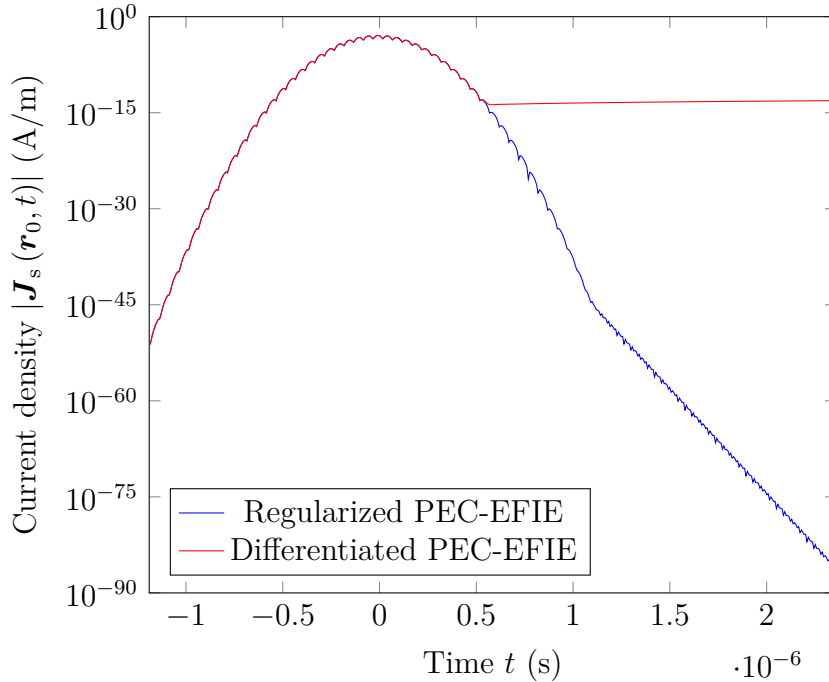


Figure 5.2: Surface current density on the unit sphere

time steps between $t_0 = -15\tau_{\text{ch}}$ and $t_{N_t} = 30\tau_{\text{ch}}$ where $\tau_{\text{ch}} = 79.6$ ns ($\Delta t = 7.16$ ns). The central frequency is $f_0 = 10$ MHz. The surface current density is computed at the position $\mathbf{r}_0 = (0.30$ m, 0.95 m, 0.06 m) (see figure 5.1a). It is clear that the regularized formulation yields the same result as the differentiated PEC-EFIE but it does not suffer from the DC instability. The DC instability is visible as a quasi constant current in the late time steps. In fact, the DC current verifies the recurrence relation in (5.34). Note that at some point, there is a saturation in the preconditioned formulation. In fact, it can be explained the same way as the DC instability by considering the eigenvalues of the companion matrix \mathbf{Z}_C (4.97). By noting $|\lambda_{\text{max}}|$ the maximum absolute value of the eigenvalues of \mathbf{Z}_C , the current at the time step j is roughly proportional to $|\lambda_{\text{max}}|^j$ in the late time steps where the excitation field has vanished. For the differentiated PEC-EFIE, $|\lambda_{\text{max}}| \approx 1$ so it results in an almost constant current, and in the regularized formulation $|\lambda_{\text{max}}| < 1$ so the current decays exponentially.

To demonstrate the applicability of the technique to multiply connected geometries, the second example in the figure 5.3 shows similar results for a torus whose inner radius is 0.6 m, outer radius is 1 m, $N_s = 1200$ and the current probe is in $\mathbf{r}_0 = (0.89$ m, 0.39 m, 0.04 m) (see figure 5.1b).

The DC instability can also be observed by plotting in the complex plane the eigenvalues of the companion matrix \mathbf{Z}_C (4.97) (i.e. the polynomial eigenvalues). In the figure 5.4, the polynomial eigenvalues associated to the matrices $\mathbf{Z}_i^{\text{der}}$ and $\mathbf{Z}_i^{\text{reg}}$ ($i \in [0, 17]$) have been plotted for a unit sphere with the parameters $N_s = 120$, $N_{\text{conv}} = 17$, $\Delta t = 7.16$ ns (it is the same set of parameters used in the section 4.4). It

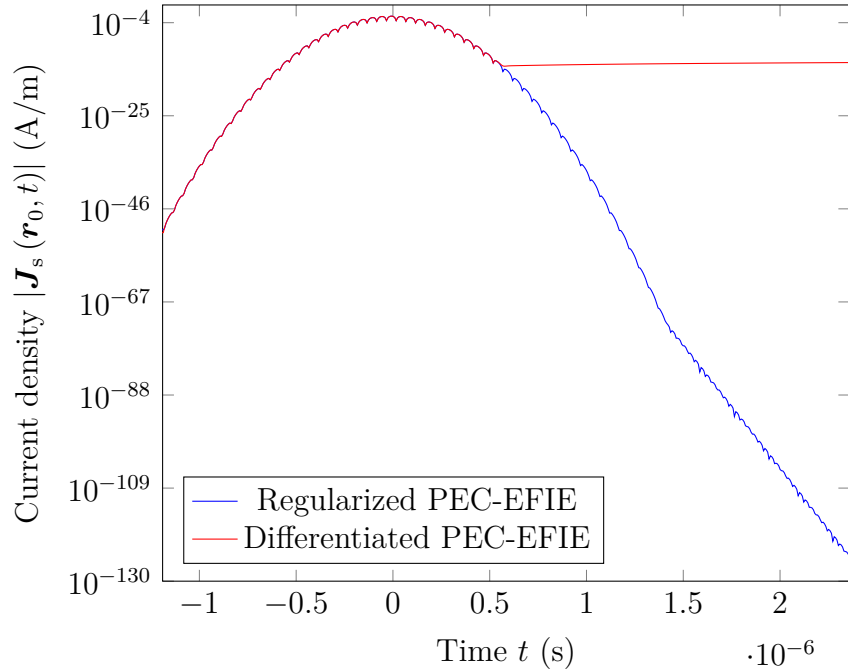


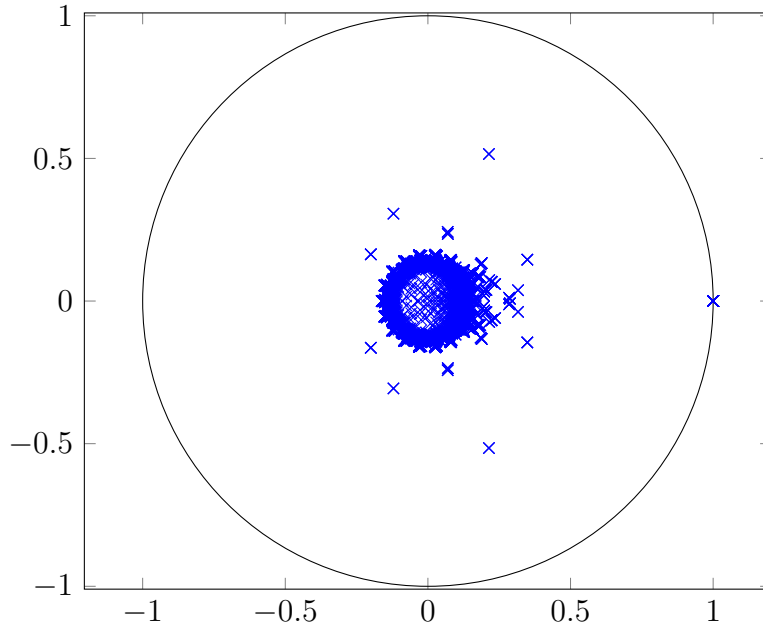
Figure 5.3: Surface current density on a torus

is clear that there is a cluster of eigenvalues in 1 for the differentiated PEC-EFIE (a zoom of it is in the section 4.4 in the figure 4.4b) which results in the DC instability. However, this cluster is removed in the regularized PEC-EFIE so it does not suffer from the DC instability. The number of matrices used in the computation of the polynomial eigenvalues is equal to $N_{\text{conv}} + 1$ (N_{conv} matrices in the convolution of the RHS plus the matrix on the LHS). In the figure 5.5, the norms of the matrices $\mathbf{Z}_i^{\text{der}}$ and $\mathbf{Z}_i^{\text{reg}}$ have been computed for $i \in [1, Q]$. It can be seen that $\|\mathbf{Z}_i^{\text{der}}\|$ and $\|\mathbf{Z}_i^{\text{reg}}\|$ decrease exponentially with the index i in the sequence. In fact, after a certain index, the matrices computed in the sequences are essentially wrong since their norms no longer decrease. In the example of the figure 5.5, it corresponds to the index $i = N_{\text{conv}} = 17$.

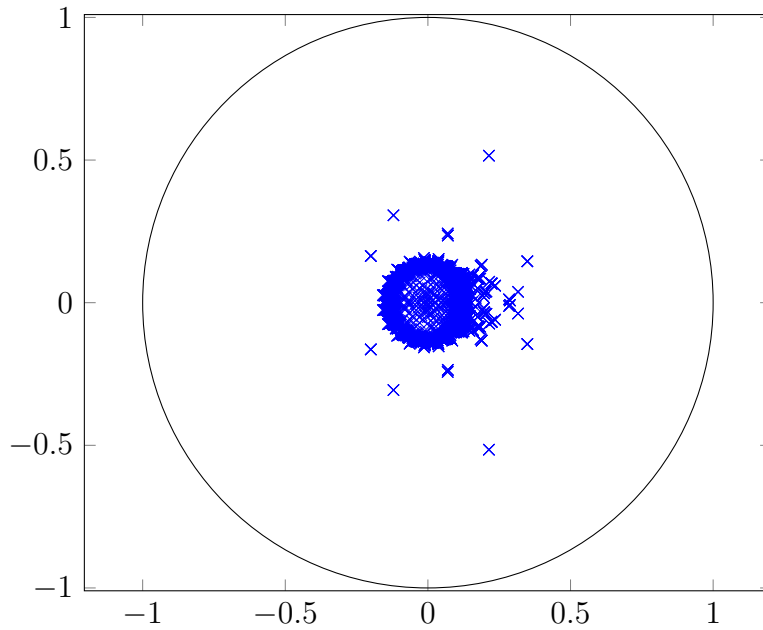
5.6.2 Results related to the large time step breakdown

The large time step stabilization is verified by showing that the condition number of the system on the LHS of the MOT is stable when the time step Δt increases. This is represented in the figure 5.6, for the unit sphere ($N_s = 1920$, $h \approx 0.15$ m). As it can be observed, the condition number of the regularized system $\text{cond}(\mathbf{Z}_i^{\text{reg}})$ tends to a constant as $\Delta t \rightarrow +\infty$ whereas the condition number of the differentiated PEC-EFIE system $\text{cond}(\mathbf{Z}_i^{\text{der}})$ scales proportionally to Δt^2 as explained in the section 5.3.1.

5. Large time step and DC stable time domain PEC-EFIE discretized with implicit Runge-Kutta methods



(a) Polynomial eigenvalues associated to $(\mathbf{Z}_i^{\text{der}})_{i=0}^{17}$ (differentiated PEC-EFIE) where there is a cluster in 1



(b) Polynomial eigenvalues associated to $(\mathbf{Z}_i^{\text{reg}})_{i=0}^{17}$ (regularized PEC-EFIE) where there is no cluster in 1

Figure 5.4: Polynomial eigenvalues associated to the sequence of interaction matrices in the differentiated PEC-EFIE and the regularized PEC-EFIE

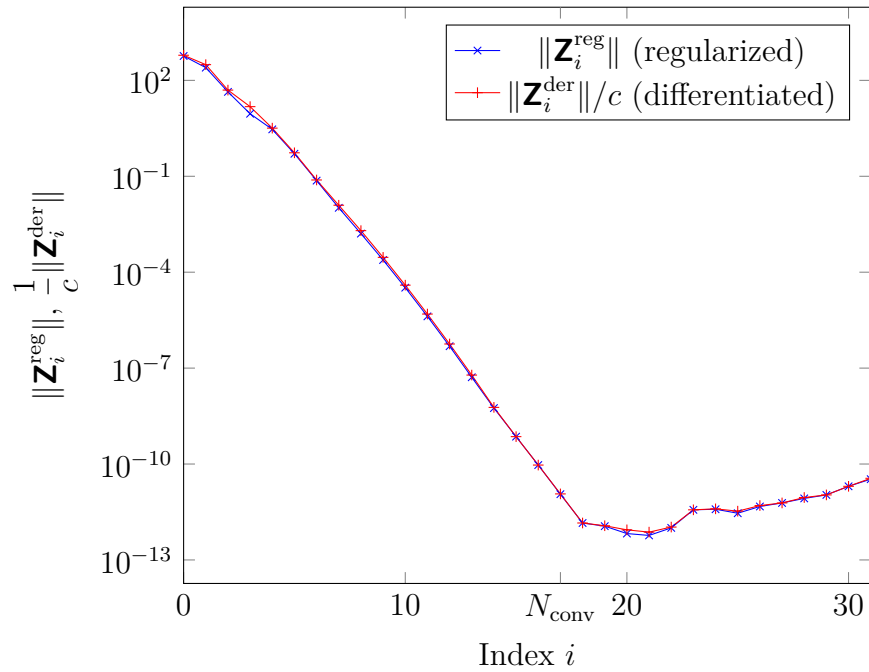


Figure 5.5: Norm of the interaction matrices of the differentiated PEC-EFIE and regularized PEC-EFIE

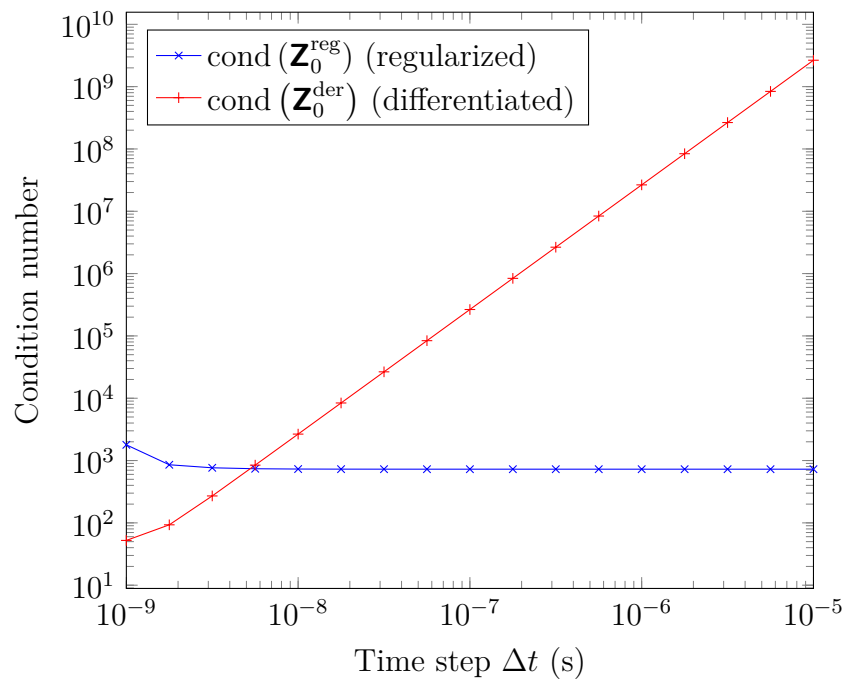


Figure 5.6: Condition number of the system solved in the MOT of the differentiated PEC-EFIE and regularized PEC-EFIE

5.7 Conclusion and future work

This chapter introduced a time domain PEC-EFIE that is stable for large time steps and does not suffer from the DC instability. The quasi-Helmholtz projectors are leveraged to obtain the regularized formulation. The projectors enables the solenoidal and non-solenoidal components to be separated and rescaled independently in the Laplace domain, which yields a stable system. Then, it has been shown that the use of the projectors is fully compatible with the IRK convolution quadrature and existing fast solvers.

Future works related to this formulation include the following topics

- Dense mesh breakdown. The formulation can be further regularized for the dense mesh breakdown by applying a Calderón preconditioning on top of the regularized MOT scheme (5.54). In fact, it was deliberately not included in this chapter to make it more concise. Essentially the goal of this chapter was to introduce the quasi-Helmholtz projectors in the IRK convolution quadrature scheme. The dense mesh stabilization is in fact simpler than the large time step stabilization because it is only a left preconditioner while the regularization presented in this chapter also requires a right preconditioner.
- Small time steps (high frequency). It must be noted that the procedure described in this chapter is effective because, the number of interaction matrices in the convolution N_{conv} on the RHS of the MOT and the number of subintervals Q required in the computation of the inverse Z-transform (4.38b) actually decrease as the time step Δt increases. Conversely, N_{conv} and Q effectively increase when the time step Δt decreases, and therefore at high frequency a strategy compatible with e.g. the Plane Wave Time Domain (PWTD) algorithm [44, 45] should be investigated.
- PEC-CFIE. A combination of the proposed PEC-EFIE formulation with its PEC-MFIE analogue results in a PEC-CFIE whose properties will be investigated. It is expected to improve the stability of the formulation at high frequency by removing the spurious resonances.
- IBC-EFIE. Since the quasi-Helmholtz projectors have been proven compatible with the PEC-EFIE, a time domain analogue of the IBC-EFIE presented in the chapter 2 is accessible. In the next chapter, a large time step and dense mesh stable IBC-EFIE is presented. The fact that the projectors with the IRK convolution quadrature have been introduced in this chapter will enable to focus on issues more specific to the IBC-EFIE.

Chapter 6

Large time step and dense mesh stable time domain IBC-EFIE

In this chapter, a formulation for the Impedance Boundary Condition - Electric Field Integral Equation (IBC-EFIE) is constructed in the time domain. The spacial discretization of the currents is achieved with a mixed discretization using Rao-Wilton-Glisson (RWG) functions for the electric current and Buffa-Christiansen (BC) functions for the magnetic current. The temporal discretization is achieved with a convolution quadrature based on Implicit Runge-Kutta (IRK) methods. First, contrary to the PEC-EFIE, the IBC-EFIE operator is lossy so the norm of the interaction matrices in the MOT does not decay exponentially fast, and therefore all the terms in the convolution on the RHS of the MOT should be computed to keep track of the losses. It is shown in this chapter how to take advantage of the lossless PEC-EFIE and PEC-MFIE operators that appear in the IBC-EFIE to efficiently compute the RHS of the MOT. And second, at each time step, the system to solve in the MOT is ill-conditioned as it suffers from the large time step and the dense mesh breakdowns. By taking advantage of the preconditioners developed in the chapter 2 in addition to the introduction of the quasi-Helmholtz projectors in the IRK convolution quadrature in the chapter 5, a stable MOT is constructed for the time domain IBC-EFIE.

6.1 Introduction

The time domain PEC-EFIE enables the analysis of a transient scattering by an object when there are no losses. As it has been explained in the sections 1.4 and 2.1, the Impedance Boundary Conditions (IBC) are widespread in the frequency domain to simulate a wide range of lossy scatterers. Indeed, in the frequency domain the impedance z^{imp} is usually a simple multiplicative factor, even when the impedance depends on the frequency. However, it is more complicated in the time domain. In general, the product with a frequency dependent impedance becomes a convolution with the impedance that is a time dependent operator when the problem is transformed into the time domain. Therefore, the impedance in the time domain usually involves several orders of higher derivatives of the fields [71]. In fact, the model for

the impedances in the time domain are usually obtained by inverse transforming the frequency or Laplace domain representation of the impedance into the time domain. If a different model for the impedance is chosen in the frequency domain, the corresponding time domain formulation is likely to change significantly. As it has been explained in the section 4.3 and illustrated in the previous chapter, the convolution quadrature do not require the expression of the convolution kernel in the time domain but its expression in the Laplace domain. In particular, this feature is used in this chapter to build a formulation for the IBC in the time domain by avoiding the need of an explicit expression for the time dependent impedance.

There exist various formulations that solve the IBC in the time domain in the context of the Boundary Element Method (BEM) [72–74]. A time discretization using convolution quadrature has been applied recently to IBC in acoustics [75, 76], but so far it seems that it has not been applied to IBC in electromagnetics. In this chapter, a formulation for the time domain IBC-EFIE is presented.

Furthermore, since the IBC introduces losses, it will be shown that a straightforward time discretization of the IBC-EFIE operator results in a MOT that involves a sequence of interaction matrices whose norms do not decay exponentially. Therefore, it requires the computation of all the terms in the convolution on the RHS of the MOT, which is prohibitively expensive computationally ($\mathcal{O}(N_t^2)$) if the number of time steps is large. Consequently, the first objective in this chapter is to reformulate the MOT of the IBC-EFIE to take advantage of the PEC-EFIE and PEC-MFIE operators that are present in the IBC-EFIE and whose sequences of interaction matrices in the time domain decay exponentially.

In addition, it is shown that the system to solve in the MOT is ill-conditioned for large time step and dense meshes. In fact, they are the analogous low frequency breakdown and dense mesh breakdown of the IBC-EFIE in the frequency domain that were addressed in the chapter 2. Thus, it is investigated how the preconditioner introduced in the chapter 2 can be used in the time domain. In particular, it makes use of quasi-Helmholtz projectors, whose compatibility with the IRK convolution quadrature has been demonstrated in the previous chapter.

This chapter is subdivided in four sections. First, the background and notations are introduced, in particular the mixed discretization in space and the IRK convolution quadrature for the discretization in time are recalled. Then, the construction of an efficient MOT is explained. Next, a preconditioner is built for the time domain IBC-EFIE based on the work of the chapters 2 (frequency domain IBC-EFIE) and 5 (time domain PEC-EFIE). Finally, numerical results are presented.

6.2 Background and notations

6.2.1 Integral equations in the time and Laplace domains

The goal is to solve the following boundary problem. An object with a boundary Γ is placed in a background medium whose permittivity is ε , permeability is μ and characteristic impedance is $\eta = \sqrt{\mu/\varepsilon}$. An incident field $\mathbf{E}^{\text{inc}}(\mathbf{r}, t)$ induces the equiv-

alent surface electric current density $\mathbf{J}_s(\mathbf{r}, t)$ and surface magnetic current density $\mathbf{M}_s(\mathbf{r}, t)$ on Γ that verify the Electric Field Integral Equation (EFIE) (see (4.14a))

$$\eta \mathcal{T} \mathbf{J}_s(\mathbf{r}, t) + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{M}_s(\mathbf{r}, t) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, t) \quad (6.1)$$

where the time domain surface integral operators \mathcal{T} , \mathcal{K} and \mathcal{I} are defined in (4.15), and $\hat{\mathbf{n}}$ is the normal to Γ . For all $t < t_0$, the fields are assumed to vanish in the neighbourhood of Γ .

In the Laplace domain, it is assumed that the currents $\mathbf{J}_s(\mathbf{r}, s)$ and $\mathbf{M}_s(\mathbf{r}, s)$ verify the following Leontovich Impedance Boundary Condition (IBC)

$$\mathbf{M}_s(\mathbf{r}, s) = -z^{\text{imp}}(s) \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{J}_s(\mathbf{r}, s) \quad (6.2)$$

where $z^{\text{imp}}(s)$ is the impedance in the Laplace domain. The corresponding relation in the time domain is a convolution in the form

$$\mathbf{M}_s(\mathbf{r}, t) = - \int_{u=0}^t \mathcal{L}^{-1}(z^{\text{imp}})(u) \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{J}_s(\mathbf{r}, t - u) du. \quad (6.3)$$

The time domain equation is discretized with the IRK convolution quadrature following the procedure described in section 4.3. So the equation must be expressed in the Laplace domain and discretized in space, then the Laplace variable s is substituted by a matrix valued $\mathbf{s}(z)$ to represent the Z-transform of the discrete time domain equation, and finally an inverse Z-transform yields a fully discretized system.

In the Laplace domain, the EFIE has an operatorial form that is similar to the one in the time domain

$$\eta \mathcal{T} \mathbf{J}_s(\mathbf{r}, s) + \left(\frac{\mathcal{I}}{2} - \mathcal{K} \right) \mathbf{M}_s(\mathbf{r}, s) = -\hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{E}^{\text{inc}}(\mathbf{r}, s) \quad (6.4)$$

where the Laplace domain integral operators \mathcal{T} and \mathcal{K} are defined in (4.83).

6.2.2 Spacial discretization

The boundary Γ is discretized into a mesh that contains N_f triangular faces, N_v vertices and N_s edges. The electric current density is discretized with the RWG basis functions $(\mathbf{f}_n)_{n=1}^{N_s}$ (defined in section 1.3.3.1), so

$$\mathbf{J}_s(\mathbf{r}, s) \approx \sum_{n=1}^{N_s} [\mathbf{J}(s)]_n \mathbf{f}_n(\mathbf{r}) \quad (6.5)$$

where $\mathbf{J}(s) \in \mathbb{C}^{N_s}$ is a vector of coefficients that represents the electric current density in the RWG basis. The magnetic current density is discretized with the BC basis functions $(\mathbf{g}_n)_{n=1}^{N_s}$ (defined in section 1.3.3.2), so

$$\mathbf{M}_s(\mathbf{r}, s) \approx \sum_{n=1}^{N_s} [\mathbf{M}(s)]_n \mathbf{g}_n(\mathbf{r}) \quad (6.6)$$

where $\mathbf{M}(s) \in \mathbb{C}^{N_s}$ is a vector of coefficients that represents the magnetic current density in the BC basis.

Both the EFIE (6.4) and the IBC (6.2) are tested with rotated RWG basis functions $\hat{\mathbf{n}} \times \mathbf{f}_m$ for each $m \in [1, N_s]$. This effectively results in a conforming system of $2N_s$ equations

$$\eta \mathbf{T}(s) \mathbf{J}(s) + \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{M}(s) = \mathbf{E}(s) \quad (6.7a)$$

$$\mathbf{G}_m \mathbf{M}(s) = -z^{\text{imp}}(s) \mathbf{G} \mathbf{J}(s) \quad (6.7b)$$

whose $2N_s$ unknowns are the coefficients of $\mathbf{J}(s)$ and $\mathbf{M}(s)$. The MoM matrices and vectors elements are defined by the inner products

$$[\mathbf{T}(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T} \mathbf{f}_n(s) \rangle \quad (6.8a)$$

$$[\mathbf{K}(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{K} \mathbf{g}_n(s) \rangle \quad (6.8b)$$

$$[\mathbf{G}_m]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathbf{g}_n \rangle \quad (6.8c)$$

$$[\mathbf{G}]_{mn} = \langle \mathbf{f}_m, \mathbf{f}_n \rangle \quad (6.8d)$$

$$[\mathbf{E}]_m = -\langle \mathbf{f}_m, \mathbf{E}^{\text{inc}}(s) \rangle \quad (6.8e)$$

that can be computed as described in the appendix A. Finally, the IBC-EFIE system matrix in the Laplace domain, which is obtained by substituting $\mathbf{M}(s)$ in (6.7a), is noted $\mathbf{S}(s)$ and defined as

$$\mathbf{S}(s) = \eta \mathbf{T}(s) - z^{\text{imp}}(s) \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{G}_m^{-1} \mathbf{G}. \quad (6.9)$$

6.2.3 Time discretization

The time discretization is achieved with a convolution quadrature based on Implicit Runge-Kutta (IRK) methods as described in the section 4.3. Hence, the Laplace variable s is first substituted by a matrix valued $\mathbf{s}(z)$ using an IRK method which yields a system in the Z-domain. Then, an inverse Z-transform is used to obtain a system in the discrete time domain. The Butcher Tableau [46]

$$\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^\top \end{array} \quad (6.10)$$

describes the p stages IRK method to be used in the discretization. The simulation is divided into N_t time steps, each with a duration Δt , starting from t_0 . The time step i starts at the time $t_i = t_0 + i\Delta t$.

6.2.3.1 Discretization of the vectors

The discretizations of the vectors \mathbf{E} , \mathbf{J} and \mathbf{M} are given in the time domain as they are evaluated at each time step $i \in [0, N_t]$ and each stage $k \in [1, p]$ at the time $t = t_i + \mathbf{c}_k \Delta t$.

According to (4.75a), the currents \mathbf{J}_s and \mathbf{M}_s are discretized in time by $(\mathbf{J}_i)_{i=0}^{N_t}$ and $(\mathbf{M}_i)_{i=0}^{N_t}$ that are the sequences of vectors in \mathbb{R}^{pN_s} that verify

$$\mathbf{J}_s(\mathbf{r}, t_i + \mathbf{c}_k \Delta t) \approx \sum_{n=1}^{N_s} [\mathbf{J}_i]_{k+(n-1)p} \mathbf{f}_n(\mathbf{r}) \quad (6.11a)$$

$$\mathbf{M}_s(\mathbf{r}, t_i + \mathbf{c}_k \Delta t) \approx \sum_{n=1}^{N_s} [\mathbf{M}_i]_{k+(n-1)p} \mathbf{g}_n(\mathbf{r}). \quad (6.11b)$$

These sequences $(\mathbf{J}_i)_i$ and $(\mathbf{M}_i)_i$ are the unknowns of the problem that need to be found.

Similarly, according to (4.76), the excitation vector is discretized as a sequence $(\mathbf{E}_i)_{i=0}^{N_t}$ computed as

$$[\mathbf{E}_i]_{k+(m-1)p} = - \langle \mathbf{f}_m, \mathbf{E}^{\text{inc}}(t_i + \mathbf{c}_k \Delta t) \rangle \quad (6.12a)$$

$$= - \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \mathbf{E}^{\text{inc}}(\mathbf{r}, t_i + \mathbf{c}_k \Delta t) \, dS. \quad (6.12b)$$

6.2.3.2 Discretization of the operators

The discretization of the operators in (6.7) is achieved with the substitution of the Laplace variable s by the matrix valued $\mathbf{s}(z)$ (4.55) that is function of the Z-domain variable z and that reads

$$\mathbf{s}(z) = \frac{1}{\Delta t} \left(\mathbf{A} + \frac{\mathbf{1b}^T}{z-1} \right)^{-1}. \quad (6.13)$$

There are three types of quantities in (6.7) that transform differently with the substitution of s by \mathbf{s} :

- The constant matrices \mathbf{G}_m and \mathbf{G} . As explained in the section 4.3.5.3, they are augmented using a Kronecker product with the $p \times p$ identity matrix \mathbf{I}_p . They are noted

$$\tilde{\mathbf{G}}_m = \mathbf{G}_m \otimes \mathbf{I}_p \quad (6.14a)$$

$$\tilde{\mathbf{G}} = \mathbf{G} \otimes \mathbf{I}_p. \quad (6.14b)$$

- The s dependent scalar impedance $z^{\text{imp}}(s)$. It transforms into $z^{\text{imp}}(\mathbf{s}(z)) \in \mathbb{C}^{p \times p}$. In fact, as explained in the section 4.3.5.3, the product of a scalar $z^{\text{imp}}(s)$ by a vector or a matrix is element-wise, so the product of a matrix valued $z^{\text{imp}}(\mathbf{s}(z))$ must be made block-wise. The block-wise multiplication by $z^{\text{imp}}(\mathbf{s}(z))$ is obtained with a block diagonal matrix $\tilde{\mathbf{z}}^{\text{imp}}(\mathbf{s}(z))$ defined as

$$\tilde{\mathbf{z}}^{\text{imp}}(\mathbf{s}(z)) = \mathbf{I} \otimes z^{\text{imp}}(\mathbf{s}(z)). \quad (6.15)$$

- The interaction matrices $\mathbf{T}(s)$ and $\mathbf{K}(s)$. They transform into the matrices $\mathbf{T}(\mathbf{s}(z))$ and $\mathbf{K}(\mathbf{s}(z))$ as explained in the section 4.3.5.2.

The z dependent quantities are transformed into sequences in the discrete time domain using the inverse \mathcal{Z} -transform. The sequence of interaction matrices $(\mathbf{T}_i)_i$ and $(\mathbf{K}_i)_i$ are defined as

$$\mathbf{T}_i = \mathcal{Z}^{-1}(\mathbf{T}(\mathbf{s}))_i \quad (6.16a)$$

$$\mathbf{K}_i = \mathcal{Z}^{-1}(\mathbf{K}(\mathbf{s}))_i. \quad (6.16b)$$

The impedance $\tilde{\mathbf{z}}^{\text{imp}}(\mathbf{s}(z))$ is also transformed into a sequence $(\tilde{\mathbf{z}}_i^{\text{imp}})_i$ defined as

$$\tilde{\mathbf{z}}_i^{\text{imp}} = \mathcal{Z}^{-1}(\tilde{\mathbf{z}}^{\text{imp}}(\mathbf{s}))_i \quad (6.17a)$$

$$= \mathbf{I} \otimes \mathcal{Z}^{-1}(z^{\text{imp}}(\mathbf{s}))_i \quad (6.17b)$$

$$= \mathbf{I} \otimes \mathbf{z}_i^{\text{imp}} \quad (6.17c)$$

$$\mathbf{z}_i^{\text{imp}} = \mathcal{Z}^{-1}(z^{\text{imp}}(\mathbf{s}))_i. \quad (6.17d)$$

Similarly, to the impedance, s is a scalar so the same kind of transformation applies. Using identical notations, the following sequence is defined

$$\tilde{\mathbf{s}}_i = \mathcal{Z}^{-1}(\tilde{\mathbf{s}})_i \quad (6.18a)$$

$$= \mathbf{I} \otimes \mathcal{Z}^{-1}(\mathbf{s})_i \quad (6.18b)$$

$$= \frac{1}{\Delta t} \left(\tilde{\mathbf{A}}^{-1} \delta_{i,0} - \tilde{\mathbf{A}}^{-1} \tilde{\mathbf{1}} \tilde{\mathbf{b}}^T \tilde{\mathbf{A}}^{-1} \delta_{i-1,0} \right) \quad (6.18c)$$

where $\tilde{\mathbf{A}}^{-1} = \mathbf{I} \otimes \mathbf{A}^{-1}$ and $\tilde{\mathbf{1}} \tilde{\mathbf{b}}^T = \mathbf{I} \otimes \mathbf{1} \mathbf{b}^T$. Note that (6.18c) is obtained with the L-stability of the IRK methods that implies (5.29). Therefore,

$$\tilde{\mathbf{s}}_0 = \frac{1}{\Delta t} \tilde{\mathbf{A}}^{-1} \quad (6.19a)$$

$$\tilde{\mathbf{s}}_1 = -\frac{1}{\Delta t} \tilde{\mathbf{A}}^{-1} \tilde{\mathbf{1}} \tilde{\mathbf{b}}^T \tilde{\mathbf{A}}^{-1}. \quad (6.19b)$$

6.2.3.3 Discretization of the equations

Each product in the Laplace domain is transformed into a convolution in the time domain that is discretized with the IRK convolution quadrature as explained in the section 4.3.4. Using the previous definitions and notations, the fully discretized time domain equations that correspond to the space discretized Laplace domain equations in (6.7) are for each time step i

$$\eta \sum_{j=0}^i \mathbf{T}_j \mathbf{J}_{i-j} + \frac{1}{2} \tilde{\mathbf{G}}_m \mathbf{M}_i - \sum_{j=0}^i \mathbf{K}_j \mathbf{M}_{i-j} = \mathbf{E}_i \quad (6.20a)$$

$$\tilde{\mathbf{G}}_m \mathbf{M}_i = -\tilde{\mathbf{G}} \sum_{j=0}^i \tilde{\mathbf{z}}_j^{\text{imp}} \mathbf{J}_{i-j}. \quad (6.20b)$$

As it is explained in the previous chapters, the time domain EFIE is usually time differentiated to be solved efficiently. This enables the norms of the interaction matrices that discretize the operator $\frac{\partial}{\partial t} \mathcal{T}$ to decay exponentially. Otherwise, the norms

of the matrices \mathbf{T}_j do not decay because of the time integration of the hypersingular part of the operator \mathcal{T} . A time derivative corresponds to a multiplication by s in the Laplace domain, so the sequences of interaction matrices that discretize the operators $\frac{\partial}{\partial t}\mathcal{T}$ and $\frac{\partial}{\partial t}\mathcal{K}$ are

$$\mathbf{T}_i^{\text{der}} = \mathcal{Z}^{-1}(\tilde{\mathbf{s}}\mathbf{T}(\mathbf{s}))_i \quad (6.21a)$$

$$\mathbf{K}_i^{\text{der}} = \mathcal{Z}^{-1}(\tilde{\mathbf{s}}\mathbf{K}(\mathbf{s}))_i. \quad (6.21b)$$

Also, the sequence of excitation vectors in the differentiated EFIE, noted $(\mathbf{E}_i^{\text{der}})_i$, is computed similarly to $(\mathbf{E}_i)_i$ (6.12) but using the time derivative of the incident field, such that

$$[\mathbf{E}_i^{\text{der}}]_{k+(m-1)p} = - \left\langle \mathbf{f}_m, \frac{\partial \mathbf{E}^{\text{inc}}}{\partial t}(t_i + \mathbf{c}_k \Delta t) \right\rangle. \quad (6.22)$$

Then, the discretization of the differentiated IBC-EFIE is

$$\eta \sum_{j=0}^i \mathbf{T}_j^{\text{der}} \mathbf{J}_{i-j} + \sum_{j=0}^i \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \right) \mathbf{M}_{i-j} = \mathbf{E}_i^{\text{der}} \quad (6.23a)$$

$$\tilde{\mathbf{G}}_m \mathbf{M}_i = -\tilde{\mathbf{G}} \sum_{j=0}^i \tilde{\mathbf{z}}_j^{\text{imp}} \mathbf{J}_{i-j}. \quad (6.23b)$$

6.3 Solution of the time differentiated IBC-EFIE

In this section the solution of the (non-preconditioned) time differentiated IBC-EFIE is explained. In a first part it is explained why an efficient Marching-On-in-Time (MOT) scheme cannot be obtained straightforwardly from (6.20) or (6.23) as it leads to lossy interaction matrices. Consequently, that results in convolutions on the RHS that require the computation of all of their terms. In a second part, a method is proposed to modify the MOT by taking advantage of the lossless PEC-EFIE and PEC-MFIE convolution kernels that are decaying exponentially fast. Therefore, they only require the computation of a fixed number of terms in the convolutions on the RHS. The third part details the computation of the auxiliary variable current that is introduced in the second part. In fact, this technique using an auxiliary variable to compute the MOT efficiently is important to understand in the case of the non-preconditioned IBC-EFIE because it is further used in the preconditioned version of the equation.

6.3.1 Overview of the problems

To obtain a MOT scheme from (6.20) or (6.23), two approaches can be immediately thought of. The first approach is to use both \mathbf{J}_i and \mathbf{M}_i as unknowns at the time step i . And the second approach is to use the IBC to substitute \mathbf{M}_i in the EFIE and solve for \mathbf{J}_i . Once \mathbf{J}_i is known, \mathbf{M}_i is computed using the IBC. This is actually similar to what is done in the frequency domain (see chapter 2). Unfortunately, these two approaches result in solvers that have a quadratic complexity in space or time as it is explained in the following sections.

6.3.1.1 Solving for the two currents simultaneously

In this section the MOT is built to be solved for a sequence of vectors containing both \mathbf{J}_i and \mathbf{M}_i at the time step i . So the two currents are solved simultaneously. The system (6.23) is rewritten in a block form as

$$\begin{pmatrix} \eta \mathbf{T}_0^{\text{der}} & \frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_0 - \mathbf{K}_0^{\text{der}} \\ \tilde{\mathbf{G}}_0^{\text{imp}} & \tilde{\mathbf{G}}_m \end{pmatrix} \begin{pmatrix} \mathbf{J}_i \\ \mathbf{M}_i \end{pmatrix} = \begin{pmatrix} \mathbf{E}_i^{\text{der}} \\ \mathbf{0} \end{pmatrix} - \sum_{j=1}^i \begin{pmatrix} \eta \mathbf{T}_j^{\text{der}} & \frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \\ \tilde{\mathbf{G}}_j^{\text{imp}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{J}_{i-j} \\ \mathbf{M}_{i-j} \end{pmatrix} \quad (6.24)$$

In general, the norm of the matrices $\tilde{\mathbf{z}}_j^{\text{imp}}$ do not decay fast enough because of the losses due to the impedance. So at the time step i , there are i terms in the convolution on the RHS that must be computed. So with N_t times steps in total, it requires $\mathcal{O}(N_t^2)$ matrix-vector products, which is unsatisfactory. There are MOT algorithms for lossy medium to reduce the number of products to $\mathcal{O}(N_t \log(N_t)^2)$ using a Fast Fourier Transform (FFT) based convolution, for example [77], but in general it also prevents from using fast matrix-vector multiplications so the complexity remains proportional to N_s^2 which is also unwanted.

6.3.1.2 Substitution of the magnetic current

Similarly to what is done in the frequency domain in the chapters 2 and 3, the magnetic current can be substituted in the EFIE by inserting the IBC (6.23b) into (6.23a). Doing so results in

$$\mathbf{M}_i = -\tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \sum_{j=0}^i \tilde{\mathbf{z}}_j^{\text{imp}} \mathbf{J}_{i-j} \quad (6.25a)$$

$$\mathbf{E}_i^{\text{der}} = \eta \sum_{j=0}^i \mathbf{T}_j^{\text{der}} \mathbf{J}_{i-j} - \sum_{j=0}^i \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \right) \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \sum_{k=0}^{i-j} \tilde{\mathbf{z}}_k^{\text{imp}} \mathbf{J}_{i-j-k} \quad (6.25b)$$

$$= \sum_{j=0}^i \left(\eta \mathbf{T}_j^{\text{der}} - \sum_{k=0}^j \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_k - \mathbf{K}_k^{\text{der}} \right) \tilde{\mathbf{z}}_{j-k}^{\text{imp}} \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \right) \mathbf{J}_{i-j}. \quad (6.25c)$$

Therefore, the MOT scheme is

$$\mathbf{S}_0^{\text{der}} \mathbf{J}_i = \mathbf{E}_i^{\text{der}} - \sum_{j=1}^i \left(\eta \mathbf{T}_j^{\text{der}} - \sum_{k=0}^j \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_k - \mathbf{K}_k^{\text{der}} \right) \tilde{\mathbf{z}}_{j-k}^{\text{imp}} \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \right) \mathbf{J}_{i-j} \quad (6.26)$$

where

$$\mathbf{S}_0^{\text{der}} = \eta \mathbf{T}_0^{\text{der}} - \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_0 - \mathbf{K}_0^{\text{der}} \right) \tilde{\mathbf{z}}_0^{\text{imp}} \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \quad (6.27a)$$

$$= \mathcal{Z}^{-1}(\tilde{\mathbf{s}}\mathbf{S}(\mathbf{s}))_0. \quad (6.27b)$$

Again, the convolution on the RHS involves a sequence of matrices whose norms decay slowly in general due to the impedance. Therefore, all the terms must be computed.

6.3.2 Marching-on-in-time

It has been explained in the previous sections that the MOT schemes in which the impedance appears in the convolutions cannot take advantage of the exponentially decaying sequences $(\mathbf{T}_j^{\text{der}})_j$ and $(\mathbf{K}_j^{\text{der}})_j$. The goal in this section is to rewrite the MOT scheme without the impedance in the convolution on the RHS. This is done by keeping the magnetic current for the computation of the RHS. Therefore, at the time step i , it is assumed that the unknowns \mathbf{J}_j and \mathbf{M}_j are known for all $j < i$. In addition, an auxiliary variable is needed, it is noted \mathbf{M}'_i and defined as

$$\mathbf{M}'_i = -\tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \sum_{j=1}^i \tilde{\mathbf{z}}_j^{\text{imp}} \mathbf{J}_{i-j} \quad (6.28a)$$

$$= \mathbf{M}_i + \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \mathbf{z}_0^{\text{imp}} \mathbf{J}_i. \quad (6.28b)$$

It is clear from its definition (6.28a) that \mathbf{M}'_i depends on all the \mathbf{J}_j for $j < i$. Therefore, it can be computed at the time step i . Then, (6.23a) is rewritten as

$$\mathbf{E}_i^{\text{der}} = \eta \sum_{j=0}^i \mathbf{T}_j^{\text{der}} \mathbf{J}_{i-j} + \sum_{j=0}^i \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \right) \mathbf{M}_{i-j} \quad (6.29a)$$

$$= \eta \sum_{j=0}^i \mathbf{T}_j^{\text{der}} \mathbf{J}_{i-j} + \sum_{j=1}^i \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \right) \mathbf{M}_{i-j} + \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_0 - \mathbf{K}_0^{\text{der}} \right) \mathbf{M}_i \quad (6.29b)$$

$$= \eta \sum_{j=1}^i \mathbf{T}_j^{\text{der}} \mathbf{J}_{i-j} + \sum_{j=1}^i \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \right) \mathbf{M}_{i-j} + \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_0 - \mathbf{K}_0^{\text{der}} \right) \mathbf{M}'_i + \mathbf{S}_0^{\text{der}} \mathbf{J}_i. \quad (6.29c)$$

So the process of solving the MOT at the time step i is the following:

1. \mathbf{M}'_i is first computed using (6.28a).
2. Then, \mathbf{J}_i is computed by solving the system

$$\mathbf{S}_0^{\text{der}} \mathbf{J}_i = \mathbf{E}_i^{\text{der}} - \eta \sum_{j=1}^i \mathbf{T}_j^{\text{der}} \mathbf{J}_{i-j} - \sum_{j=1}^i \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_j - \mathbf{K}_j^{\text{der}} \right) \mathbf{M}_{i-j} - \left(\frac{1}{2} \tilde{\mathbf{G}}_m \tilde{\mathbf{s}}_0 - \mathbf{K}_0^{\text{der}} \right) \mathbf{M}'_i. \quad (6.30)$$

3. Finally, when \mathbf{J}_i is known, \mathbf{M}_i is computed using

$$\mathbf{M}_i = \mathbf{M}'_i - \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \mathbf{z}_0^{\text{imp}} \mathbf{J}_i. \quad (6.31)$$

It is clear from (6.30) that the interaction matrices in the convolutions are the same as the differentiated PEC-EFIE and PEC-MFIE (they are present in the PEC-CFIE [68]). Therefore, they can be computed efficiently (the number of terms in the convolution can be truncated to a constant N_{conv}). It only remains to explain how \mathbf{M}'_i is computed at each time step.

6.3.3 Computation of the auxiliary current

To make the notations simpler, the following matrix is used

$$\mathbf{Z}_i = -\tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \mathbf{z}_i^{\text{imp}} \quad (6.32)$$

such that

$$\mathbf{M}'_i = \sum_{j=1}^i \mathbf{Z}_j \mathbf{J}_{i-j} \quad (6.33a)$$

$$= \mathbf{M}_i - \mathbf{Z}_0 \mathbf{J}_i. \quad (6.33b)$$

Naively, the computation of \mathbf{M}'_i requires i products for each $i \in [1, N_t]$ and each product has a complexity $\mathcal{O}(N_s)$. So overall the computational complexity is $\mathcal{O}(N_s N_t^2)$. Fortunately, there is a fast way of computing \mathbf{M}'_i that avoids the quadratic cost.

The goal is to find a fast algorithm to compute (6.33a) when \mathbf{J}_j is known for each $j < i$ (\mathbf{J}_i is not known) and \mathbf{Z}_i is known for all i . In fact, since the matrix \mathbf{Z}_i corresponds to a lossy interaction, the algorithm described in [77] can be applied as described in the following. The convolution is first rewritten as a block matrix-vector product

$$\begin{pmatrix} \mathbf{0} \\ \mathbf{Z}_1 \mathbf{0} \\ \mathbf{Z}_2 \mathbf{Z}_1 \mathbf{0} & (\mathbf{0}) \\ \mathbf{Z}_3 \mathbf{Z}_2 \mathbf{Z}_1 \mathbf{0} \\ \mathbf{Z}_4 \mathbf{Z}_3 \mathbf{Z}_2 \mathbf{Z}_1 \mathbf{0} \\ \mathbf{Z}_5 \mathbf{Z}_4 \mathbf{Z}_3 \mathbf{Z}_2 \mathbf{Z}_1 \mathbf{0} \\ \mathbf{Z}_6 \mathbf{Z}_5 \mathbf{Z}_4 \mathbf{Z}_3 \mathbf{Z}_2 \mathbf{Z}_1 \mathbf{0} \\ \mathbf{Z}_7 \mathbf{Z}_6 \mathbf{Z}_5 \mathbf{Z}_4 \mathbf{Z}_3 \mathbf{Z}_2 \mathbf{Z}_1 \mathbf{0} \\ \vdots \\ \ddots \end{pmatrix} \begin{pmatrix} \mathbf{J}_0 \\ \mathbf{J}_1 \\ \mathbf{J}_2 \\ \mathbf{J}_3 \\ \mathbf{J}_4 \\ \mathbf{J}_5 \\ \mathbf{J}_6 \\ \mathbf{J}_7 \\ \vdots \end{pmatrix} = \begin{pmatrix} \mathbf{M}'_0 \\ \mathbf{M}'_1 \\ \mathbf{M}'_2 \\ \mathbf{M}'_3 \\ \mathbf{M}'_4 \\ \mathbf{M}'_5 \\ \mathbf{M}'_6 \\ \mathbf{M}'_7 \\ \vdots \end{pmatrix} \quad (6.34)$$

The blocks in this matrix are Toeplitz matrices. Any, $N \times N$ Toeplitz matrix can be multiplied by a vector in $\mathcal{O}(N \log(N))$ operations using an FFT instead of the usual $\mathcal{O}(N^2)$ cost of a matrix-vector product. The $N \times N$ Toeplitz blocks are stored as the $2N$ eigenvalues of the corresponding circulant matrix of size $2N \times 2N$.

Because of their particular forms, the matrices \mathbf{Z}_i can be factored as a constant matrix part $\tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}}$ and a time dependent impedance $\mathbf{z}_i^{\text{imp}} \in \mathbb{C}^{p \times p}$ so that only the impedances should be computed and stored as eigenvalues. The constant matrix part that is $\tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}}$ is not stored directly. Instead, the sparse matrices \mathbf{G} and \mathbf{G}_m are stored and their product by a vector is computed in a $\mathcal{O}(N_s)$ complexity.

In one Toeplitz block, there are at $\mathcal{O}(N_t)$ impedances that are transformed into (and stored as) the eigenvalues of the corresponding circulant matrix in $\mathcal{O}(N_t \log(N_t))$ operations using an FFT. Since there are $\mathcal{O}(\log(N_t))$ different blocks, the overall setup costs $\mathcal{O}(N_t \log(N_t)^2)$.

There are overall $\mathcal{O}(N_t \log(N_t))$ element-wise multiplications between the eigenvalues and the Fourier-transformed vectors and each cost $\mathcal{O}(N_s)$, so overall the elements multiplications cost $\mathcal{O}(N_s N_t \log(N_t))$.

Finally, at each time step i , the block that contains the element \mathbf{Z}_1 is multiplied by the corresponding block vector and added to the result. This operation completes the computation of M_i^t needed at the time step i , and it also partially computes the sums for the next elements. Overall, the FFTs and inverse FFTs require $\mathcal{O}(N_s N_t \log(N_s N_t) \log(N_t))$ operations, which is the dominant cost.

6.4 Preconditioning of the time domain IBC-EFIE

6.4.1 Overview of the problems

The matrix $\mathbf{S}_0^{\text{der}}$ on the LHS of (6.30) is ill-conditioned. Using the results from the low frequency breakdown analysis of the IBC-EFIE that is done in the section 2.3.3, an analysis similar to the one done in the section 5.3.1 for the large time step breakdown of the PEC-EFIE can be done for the IBC-EFIE. The condition number of the system in (6.30) increases proportionally to the time step Δt (assuming a non-zero impedance) which is the large time step breakdown. The condition number is also proportional to the inverse of the mesh edge length h (see section 2.3.4) which is the dense mesh breakdown.

In the chapter 2, a low frequency and dense mesh stable preconditioner is constructed for the IBC-EFIE. Using the same preconditioning strategy, it suggests the use of the following left and right preconditioners in the Laplace domain to solve the low frequency breakdown (see (2.36) and (2.37))

$$\mathbf{L}(s) = \frac{c}{sa} \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \quad (6.35a)$$

$$\mathbf{R}(s) = \frac{\frac{sa}{c}}{\frac{sa}{c} + \frac{z^{\text{imp}}(s)}{\eta}} \mathbf{P}^{\Lambda H} + \frac{sa}{c} \mathbf{P}^{\Sigma} \quad (6.35b)$$

where a is a length that is chosen to be the diameter of Γ . On top of it, a Calderón-like preconditioning is used to solve the dense mesh breakdown (see (2.46)). So, in the Laplace domain a similar preconditioner is used,

$$\tilde{\mathbf{T}}(s) = \frac{1}{a} \mathbf{P}^{\Sigma H} \mathbb{T}_s(s) \mathbf{P}^{\Sigma H} + \mathbf{P}^{\Lambda} \quad (6.36)$$

where $\mathbb{T}_s(s)$ is the discretization of the operator \mathcal{T} with BC basis functions i.e. its elements are $[\mathbb{T}_s(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{g}_m, \mathcal{T} \mathbf{g}_n(s) \rangle$.

Inserting the IBC (6.7b) in the EFIE (6.7a) results in the following IBC-EFIE system

$$\mathbf{S}(s) \mathbf{J}(s) = \mathbf{E}(s) \quad (6.37)$$

where $\mathbf{S}(s)$ is defined in (6.9). This system is preconditioned as

$$\tilde{\mathbf{T}}(s) \mathbf{G}_m^{-1} \mathbf{L}(s) \mathbf{S}(s) \mathbf{R}(s) \mathbf{Y}(s) = \tilde{\mathbf{T}}(s) \mathbf{G}_m^{-1} \mathbf{V}(s) \quad (6.38)$$

where the original unknown current $\mathbf{J}(s)$ was replaced by the auxiliary unknown $\mathbf{Y}(s)$ and the RHS $\mathbf{E}(s)$ was replaced by $\mathbf{V}(s)$ that are defined by the relations

$$\mathbf{R}(s)\mathbf{Y}(s) = \mathbf{J}(s) \quad (6.39a)$$

$$\mathbf{V}(s) = \mathbf{L}(s)\mathbf{E}(s). \quad (6.39b)$$

The challenging part is due to the preconditioner that solves the large time step (low frequency) breakdown. Especially, the right preconditioner (\mathbf{R}) and the change of unknown (\mathbf{Y} instead of \mathbf{J}) complicate the MOT:

- The preconditioning in the frequency or Laplace domains only takes the electric current \mathbf{J} into account since the magnetic current \mathbf{M} is substituted into the EFIE (the IBC-EFIE operator \mathcal{S} acts only on \mathbf{J}). However, as it is explained in the previous section, in the case of IBC in the time domain, the magnetic current must be kept in order to avoid the impedance on the RHS. The corresponding variables were \mathbf{M}_i and \mathbf{M}'_i .
- The preconditioner \mathbf{R} also has an impedance appearing in its expression. So this impedance must also be taken away from the RHS using new auxiliary variables that have roles similar to \mathbf{M}_i and \mathbf{M}'_i . These new auxiliary variables are introduced later as \mathbf{W}_i and its corresponding \mathbf{W}'_i .
- The preconditioning involves a change of unknown as the system is now solved for \mathbf{Y} instead of \mathbf{J} . So, the electric current \mathbf{J}_i also has an auxiliary variable \mathbf{J}'_i .

The preconditioner for the dense mesh breakdown is purely multiplicative so it does not introduce complications in the formulation.

6.4.2 Large time step stabilization

6.4.2.1 Left hand side

The low frequency stable matrix in the Laplace domain is named \mathbf{S}^{stab} such that $\mathbf{S}^{\text{stab}}(s) = \mathbf{L}(s)\mathbf{S}(s)\mathbf{R}(s)$. In addition, two scalar variables (no units) are introduced to simplify the expressions in the following

$$\alpha(s) = \frac{\frac{sa}{c}}{\frac{sa}{c} + \frac{z^{\text{imp}}(s)}{\eta}} \quad (6.40a)$$

$$\beta(s) = \frac{\frac{z^{\text{imp}}(s)}{\eta}}{\frac{sa}{c} + \frac{z^{\text{imp}}(s)}{\eta}}. \quad (6.40b)$$

Finally, $\mathbf{T}(s)$ is rewritten as the sum of its two parts

$$\mathbf{T}(s) = -\frac{s}{c}\mathbf{T}_s(s) + \frac{c}{s}\mathbf{T}_h(s) \quad (6.41a)$$

$$[\mathbf{T}_s(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_s \mathbf{f}_n(s) \rangle \quad (6.41b)$$

$$[\mathbf{T}_h(s)]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_h \mathbf{f}_n(s) \rangle. \quad (6.41c)$$

Then, the expansion of the product that defines $\mathbf{S}^{\text{stab}}(s)$ is

$$\mathbf{S}^{\text{stab}}(s) = \mathbf{L}(s)\mathbf{S}(s)\mathbf{R}(s) \quad (6.42a)$$

$$\begin{aligned} &= \eta \mathbf{P}^{\Lambda H} \left(-\alpha(s) \frac{1}{a} \mathbf{T}_s(s) - \beta(s) \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{G}_m^{-1} \mathbf{G} \right) \mathbf{P}^{\Lambda H} \\ &+ \eta \mathbf{P}^{\Lambda H} \left(-\frac{s}{c} \mathbf{T}_s(s) - \frac{z^{\text{imp}}(s)}{\eta} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{G}_m^{-1} \mathbf{G} \right) \mathbf{P}^{\Sigma} \end{aligned} \quad (6.42b)$$

$$\begin{aligned} &+ \eta \mathbf{P}^{\Sigma} \left(-\alpha(s) \frac{s}{c} \mathbf{T}_s(s) - \beta(s) \frac{sa}{c} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{G}_m^{-1} \mathbf{G} \right) \mathbf{P}^{\Lambda H} \\ &+ \eta \mathbf{P}^{\Sigma} \left(-\frac{s^2 a}{c^2} \mathbf{T}_s(s) + a \mathbf{T}_h(s) - \frac{z^{\text{imp}}(s)}{\eta} \frac{sa}{c} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{G}_m^{-1} \mathbf{G} \right) \mathbf{P}^{\Sigma}. \end{aligned}$$

This expansion of $\mathbf{S}^{\text{stab}}(s)$ is useful for the computation of the LHS in an iterative solver. As it is done for $z^{\text{imp}}(s)$ in (6.17), the s dependent scalars $\alpha(s)$, $\beta(s)$ and s are transformed into the discrete time domain, resulting in the sequences of block diagonal matrices

$$\tilde{\mathbf{z}}_i^{\text{imp}} = \mathbf{I} \otimes \mathcal{Z}^{-1}(z^{\text{imp}}(\mathbf{s}))_i \quad (6.43a)$$

$$\tilde{\boldsymbol{\alpha}}_i = \mathbf{I} \otimes \mathcal{Z}^{-1}(\boldsymbol{\alpha}(\mathbf{s}))_i \quad (6.43b)$$

$$\tilde{\boldsymbol{\beta}}_i = \mathbf{I} \otimes \mathcal{Z}^{-1}(\boldsymbol{\beta}(\mathbf{s}))_i \quad (6.43c)$$

$$\tilde{\mathbf{s}}_i = \mathbf{I} \otimes \mathcal{Z}^{-1}(\mathbf{s})_i. \quad (6.43d)$$

Similarly to the sequences $(\mathbf{T}_i)_i$ and $(\mathbf{K}_i)_i$ defined in (6.16), the sequences associated to the operators \mathcal{T}_s and \mathcal{T}_h are

$$\mathbf{T}_{si} = \mathcal{Z}^{-1}(\mathbf{T}_s(\mathbf{s}))_i \quad (6.44a)$$

$$\mathbf{T}_{hi} = \mathcal{Z}^{-1}(\mathbf{T}_h(\mathbf{s}))_i. \quad (6.44b)$$

And finally, the projectors are independent of s so they are transformed like \mathbf{G}_m or \mathbf{G} in (6.14), using a Kronecker product with the $p \times p$ identity \mathbf{I}_p to take into account the p stages

$$\tilde{\mathbf{P}}^{\Lambda H} = \mathbf{P}^{\Lambda H} \otimes \mathbf{I}_p \quad (6.45a)$$

$$\tilde{\mathbf{P}}^{\Sigma} = \mathbf{P}^{\Sigma} \otimes \mathbf{I}_p. \quad (6.45b)$$

With the above notations, the matrix $\mathbf{S}_0^{\text{stab}}$ that appears on the LHS of the stable MOT is

$$\mathbf{S}_0^{\text{stab}} = \mathcal{Z}^{-1}(\mathbf{S}^{\text{stab}}(\mathbf{s}))_0 \quad (6.46a)$$

$$\begin{aligned} &= \eta \tilde{\mathbf{P}}^{\Lambda H} \left(-\tilde{\alpha}_0 \frac{1}{a} \mathbf{T}_{s0} - \tilde{\beta}_0 \left(\frac{1}{2} \tilde{\mathbf{G}}_m - \mathbf{K}_0 \right) \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \right) \tilde{\mathbf{P}}^{\Lambda H} \\ &+ \eta \tilde{\mathbf{P}}^{\Lambda H} \left(-\frac{\tilde{\mathbf{s}}_0}{c} \mathbf{T}_{s0} - \frac{\tilde{\mathbf{z}}_0^{\text{imp}}}{\eta} \left(\frac{1}{2} \tilde{\mathbf{G}}_m - \mathbf{K}_0 \right) \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \right) \tilde{\mathbf{P}}^{\Sigma} \end{aligned} \quad (6.46b)$$

$$\begin{aligned} &+ \eta \tilde{\mathbf{P}}^{\Sigma} \left(-\tilde{\alpha}_0 \frac{\tilde{\mathbf{s}}_0}{c} \mathbf{T}_{s0} - \tilde{\beta}_0 \frac{\tilde{\mathbf{s}}_0 a}{c} \left(\frac{1}{2} \tilde{\mathbf{G}}_m - \mathbf{K}_0 \right) \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \right) \tilde{\mathbf{P}}^{\Lambda H} \\ &+ \eta \tilde{\mathbf{P}}^{\Sigma} \left(-\frac{\tilde{\mathbf{s}}_0^2 a}{c^2} \mathbf{T}_{s0} + a \mathbf{T}_{h0} - \frac{\tilde{\mathbf{z}}_0^{\text{imp}} \tilde{\mathbf{s}}_0 a}{\eta c} \left(\frac{1}{2} \tilde{\mathbf{G}}_m - \mathbf{K}_0 \right) \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \right) \tilde{\mathbf{P}}^{\Sigma}. \end{aligned}$$

By taking into account the remarks done in the previous chapter in the section 5.5, each matrix with a tilde in this expression can be multiplied by a vector in a $\mathcal{O}(N_s)$ complexity, and the matrices without tilde are computed with a $\mathcal{O}(QN_s \log(N_s))$ complexity if a fast solver is used. Therefore, the multiplication of $\mathbf{S}_0^{\text{stab}}$ by a vector can be done in a $\mathcal{O}(QN_s \log(N_s))$ complexity.

6.4.2.2 Right hand side

The preconditioned excitation vector (6.39b) is discretized the same way as in the previous chapter as

$$\mathbf{V}_i = \frac{c}{a} \tilde{\mathbf{P}}^{\Lambda H} \mathbf{E}_i^{\text{prim}} + \tilde{\mathbf{P}}^{\Sigma} \mathbf{E}_i \quad (6.47)$$

where $(\mathbf{E}_i^{\text{prim}})_i$ is a sequence that discretize a primitive of the incident plane wave \mathbf{E}^{prim} such that $\frac{\partial \mathbf{E}^{\text{prim}}}{\partial t} = \mathbf{E}^{\text{inc}}$. It can be computed analytically as explained in the section 5.4.2.2 or numerically as explained in the section 5.4.2.3.

To build an efficient MOT, the product $\mathbf{S}^{\text{stab}}(s)\mathbf{Y}(s)$ is rewritten such that $z^{\text{imp}}(s)$ and $1/s$ do not appear in the expression. Indeed, the presence of these terms in the time domain results in non-exponentially decaying sequences in the convolutions on the RHS, which in turn implies an number of terms to compute that is not bounded in these convolutions. To achieve the removal of $z^{\text{imp}}(s)$ and $1/s$, the following variable is introduced

$$\mathbf{W}(s) = \frac{c}{sa} \mathbf{M}(s) \quad (6.48a)$$

$$= -\frac{c}{sa} z^{\text{imp}}(s) \mathbf{G}_m^{-1} \mathbf{G} \mathbf{J}(s) \quad (6.48b)$$

$$= -\eta \beta(s) \mathbf{G}_m^{-1} \mathbf{G} \mathbf{P}^{\Lambda H} \mathbf{Y}(s) - z^{\text{imp}}(s) \mathbf{G}_m^{-1} \mathbf{G} \mathbf{P}^{\Sigma} \mathbf{Y}(s). \quad (6.48c)$$

Then, $\mathbf{S}^{\text{stab}}(s)\mathbf{Y}(s)$ is rewritten as

$$\mathbf{S}^{\text{stab}}(s)\mathbf{Y}(s) = \mathbf{L}(s)\mathbf{S}(s)\mathbf{R}(s)\mathbf{Y}(s) \quad (6.49a)$$

$$= \left(\frac{c}{sa} \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \right) \left(-\eta \frac{s}{c} \mathbf{T}_s(s) \right) \mathbf{J}(s) + \eta a \mathbf{T}_h(s) \mathbf{Y}(s) \quad (6.49b)$$

$$+ \left(\frac{c}{sa} \mathbf{P}^{\Lambda H} + \mathbf{P}^{\Sigma} \right) \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{M}(s) \\ = -\eta \frac{1}{a} \mathbf{P}^{\Lambda H} \mathbf{T}_s(s) \mathbf{J}(s) - \eta \frac{s}{c} \mathbf{P}^{\Sigma} \mathbf{T}_s(s) \mathbf{J}(s) + \eta a \mathbf{T}_h(s) \mathbf{Y}(s) \quad (6.49c) \\ + \mathbf{P}^{\Lambda H} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{W}(s) + \mathbf{P}^{\Sigma} \left(\frac{1}{2} \mathbf{G}_m - \mathbf{K}(s) \right) \mathbf{M}(s).$$

At this point, the strategy to write the MOT is similar to what is done in the section (6.3.2), in the sense that without preconditioning, the unknown is \mathbf{J}_i and the auxiliary variable is the magnetic current \mathbf{M}_i and its corresponding \mathbf{M}'_i . With the preconditioning, the unknown is \mathbf{Y}_i and there are 3 auxiliary variables \mathbf{J}_i , \mathbf{M}_i , \mathbf{W}_i each with their corresponding \mathbf{J}'_i , \mathbf{M}'_i , \mathbf{W}'_i that are updated using the method described in the section 6.3.3. This is summarized in the table 6.1. There are the following equations to update the auxiliary variables \mathbf{J}_i , \mathbf{M}_i , \mathbf{W}_i when \mathbf{Y}_i has been solved at the time step i

$$\mathbf{J}_i = \mathbf{J}'_i + \left(\tilde{\boldsymbol{\alpha}}_0 \tilde{\mathbf{P}}^{\Lambda H} + \frac{\tilde{\mathbf{s}}_0 a}{c} \tilde{\mathbf{P}}^{\Sigma} \right) \mathbf{Y}_i \quad (6.50a)$$

$$\mathbf{M}_i = \mathbf{M}'_i - \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \tilde{\mathbf{z}}_0^{\text{imp}} \mathbf{J}_i \quad (6.50b)$$

$$\mathbf{W}_i = \mathbf{W}'_i - \left(\eta \tilde{\boldsymbol{\beta}}_0 \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \tilde{\mathbf{P}}^{\Lambda H} + \tilde{\mathbf{z}}_0^{\text{imp}} \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \tilde{\mathbf{P}}^{\Sigma} \right) \mathbf{Y}_i \quad (6.50c)$$

where the primed quantities are

$$\mathbf{J}'_i = \tilde{\mathbf{P}}^{\Lambda H} \sum_{j=1}^i \tilde{\boldsymbol{\alpha}}_j \mathbf{Y}_{i-j} + \tilde{\mathbf{P}}^{\Sigma} \sum_{j=1}^i \frac{\tilde{\mathbf{s}}_j a}{c} \mathbf{Y}_{i-j} \quad (6.51a)$$

$$\mathbf{M}'_i = -\tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \sum_{j=1}^i \tilde{\mathbf{z}}_j^{\text{imp}} \mathbf{J}_{i-j} \quad (6.51b)$$

$$\mathbf{W}'_i = -\eta \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \tilde{\mathbf{P}}^{\Lambda H} \sum_{j=1}^i \tilde{\boldsymbol{\beta}}_j \mathbf{Y}_{i-j} - \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \tilde{\mathbf{P}}^{\Sigma} \sum_{j=1}^i \tilde{\mathbf{z}}_j^{\text{imp}} \mathbf{Y}_{i-j}. \quad (6.51c)$$

Again, each of the above convolutions involving $\tilde{\boldsymbol{\alpha}}_j$, $\tilde{\boldsymbol{\beta}}_j$ or $\tilde{\mathbf{z}}_j^{\text{imp}}$ is computed using the technique described in 6.3.3. The convolution involving $\tilde{\mathbf{s}}_j$ is in fact a single term according to (6.18c).

Also, the sequence of matrices $(\mathbf{T}_{s_i}^{\text{der}})_i$ is defined similarly to (6.21) as

$$\mathbf{T}_{s_i}^{\text{der}} = \mathcal{Z}^{-1}(\tilde{\mathbf{s}} \mathbf{T}_s(s))_i. \quad (6.52)$$

The equation $\mathbf{S}^{\text{stab}}(s)\mathbf{Y}(s) = \mathbf{V}(s)$ is discretized using the expression (6.49c). The MOT scheme is obtained by isolating \mathbf{Y}_i on the LHS of the equation that is $\mathbf{S}_0^{\text{stab}} \mathbf{Y}_i$.

Formulation	Unknown	Auxiliary variables
Differentiated IBC-EFIE	\mathbf{J}_i	$\mathbf{M}_i, \mathbf{M}'_i$
Preconditioned IBC-EFIE	\mathbf{Y}_i	$\mathbf{J}_i, \mathbf{J}'_i, \mathbf{M}_i, \mathbf{M}'_i, \mathbf{W}_i, \mathbf{W}'_i$

Figure 6.1: Unknowns solved by the MOT and auxiliary variables used for the computation of the RHS of the differentiated (non-preconditioned) and the preconditioned IBC-EFIEs

The RHS noted $\mathbf{V}_i^{\text{stab}}$ is written such that all the convolutions have products involving exponentially decaying sequences of matrices. The MOT reads for each time step i

$$\mathbf{S}_0^{\text{stab}} \mathbf{Y}_i = \mathbf{V}_i^{\text{stab}} \quad (6.53)$$

where $\mathbf{S}_0^{\text{stab}}$ is defined in (6.46), and the RHS is

$$\mathbf{V}_i^{\text{stab}} = \mathbf{V}_i + \frac{\eta}{a} \tilde{\mathbf{P}}^{\Lambda H} \sum_{j=1}^i \mathbf{T}_{sj} \mathbf{J}_{i-j} + \frac{\eta}{c} \tilde{\mathbf{P}}^{\Sigma} \sum_{j=1}^i \mathbf{T}_{sj}^{\text{der}} \mathbf{J}_{i-j} - \eta a \sum_{j=1}^i \mathbf{T}_{hj} \mathbf{Y}_{i-j} \quad (6.54a)$$

$$+ \tilde{\mathbf{P}}^{\Lambda H} \sum_{j=1}^i \mathbf{K}_j \mathbf{W}_{i-j} + \tilde{\mathbf{P}}^{\Sigma} \sum_{j=1}^i \mathbf{K}_j \mathbf{M}_{i-j} \quad (6.54b)$$

$$+ \eta \left(\frac{1}{a} \tilde{\mathbf{P}}^{\Lambda H} + \frac{\tilde{\mathbf{S}}_0}{c} \tilde{\mathbf{P}}^{\Sigma} \right) \mathbf{T}_{s0} \mathbf{J}'_i \quad (6.54c)$$

$$- \tilde{\mathbf{P}}^{\Lambda H} \left(\frac{1}{2} \tilde{\mathbf{G}}_m - \mathbf{K}_0 \right) \mathbf{W}'_i - \tilde{\mathbf{P}}^{\Sigma} \left(\frac{1}{2} \tilde{\mathbf{G}}_m - \mathbf{K}_0 \right) \left(\mathbf{M}'_i - \tilde{\mathbf{G}}_m^{-1} \tilde{\mathbf{G}} \mathbf{z}_0^{\text{imp}} \mathbf{J}'_i \right). \quad (6.54d)$$

It must be noted that each convolution can be truncated to a fixed number of terms N_{conv} . Also, as explained in the previous chapter in the section 5.5, each quantity on this RHS can be computed efficiently using fast algorithms in a $\mathcal{O}(QN_s \log(N_s))$ complexity.

6.4.3 Dense mesh stabilization

The condition number of $\mathbf{S}_0^{\text{stab}}$ is stable for large time steps but it remains unstable when the density of the discretization increases. Fortunately, compared to the large time step preconditioning, the dense mesh preconditioning is simpler as it only involves a left preconditioning. There is no need to modify the MOT and the preconditioner can be simply applied on top of (6.53).

Regarding the dense mesh preconditioner in the Laplace domain $\tilde{\mathbf{T}}(s)$ (6.36), the matrix that preconditions $\mathbf{S}_0^{\text{stab}}$ is

$$\tilde{\mathbf{T}}_0 = \mathcal{Z}^{-1}(\tilde{\mathbf{T}}(\mathbf{s}))_0 \quad (6.55a)$$

$$= \frac{1}{a} \tilde{\mathbf{P}}^{\Sigma H} \mathbf{T}_{s0} \tilde{\mathbf{P}}^{\Sigma H} + \tilde{\mathbf{P}}^{\Lambda} \quad (6.55b)$$

where

$$\tilde{\mathbf{P}}^{\Sigma H} = \mathbf{P}^{\Sigma H} \otimes \mathbf{I}_p \quad (6.56a)$$

$$\tilde{\mathbf{P}}^\Lambda = \tilde{\mathbf{P}}^\Lambda \otimes \mathbf{I}_p \quad (6.56b)$$

$$\mathbb{T}_{s0} = \mathcal{Z}^{-1}(\mathbb{T}_s(\mathbf{s}))_0. \quad (6.56c)$$

Since this preconditioner is discretized with BC basis functions, the inverse of the mix-Gram matrix $\tilde{\mathbf{G}}_m^{-1}$ is used to apply it conformingly on the system discretized with RWG basis functions. Therefore, the MOT that is large time step and dense mesh stable reads

$$\tilde{\mathbb{T}}_0 \tilde{\mathbf{G}}_m^{-1} \mathbf{S}_0^{\text{stab}} \mathbf{Y}_i = \tilde{\mathbb{T}}_0 \tilde{\mathbf{G}}_m^{-1} \mathbf{V}_i^{\text{stab}}. \quad (6.57)$$

6.5 Numerical results

In the following results, the impedance is chosen to be proportional to the square root of the frequency as

$$z^{\text{imp}}(s) = \left(\frac{\mu_0 s}{\sigma'} \right)^{\frac{1}{2}} \quad (6.58)$$

with $\mu_0 = 4\pi 10^{-7}$ H/m and $\sigma' = 1$ S/m.

Also, the IRK method is the 3 stages Radau IIA method (4.45). In the computation of the inverse Z-transform, $Q = 32$ and $\rho = 1.0001$ have been used in the trapezoidal rules (4.38b). In the preconditioners, the length parameter a is kept to 1 m.

6.5.1 Correctness of the formulation

To verify the correctness of the formulation, the frequency domain IBC-EFIE (see chapter 2) can be used. The frequency domain impedance that corresponds to (6.58) is obtained with the substitution $s = i\omega$, and is

$$z^{\text{imp}}(\omega) = \sqrt{\frac{\mu_0 \omega}{2\sigma'}} (1 + i). \quad (6.59)$$

Assuming that the simulation is done between t_0 and t_{N_t} , the amplitude of the excitation field is sampled every Δt so that there are $N_t + 1$ samples in the time domain. Computing the FFT of the amplitude of the excitation field samples results in $N_t + 1$ samples in the frequency domain that correspond to the complex valued amplitude of plane waves (1.42). The step between each frequency sample is $\Delta f = 1/(N_t \Delta t)$ between $-f_s/2$ and $f_s/2$ where $f_s = 1/\Delta t$ is the sampling frequency. The IBC problem is then solved to obtain the solution currents for each positive frequency, and the solution current in negative frequency is obtained by taking the conjugate of the corresponding current in positive frequency. Finally, computing the inverse FFT of the sequence of frequency domain solution currents yields the time domain sequence of currents. Even though the reference solution computed this way is limited by the machine precision, it is enough to ensure the correctness of the formulation.

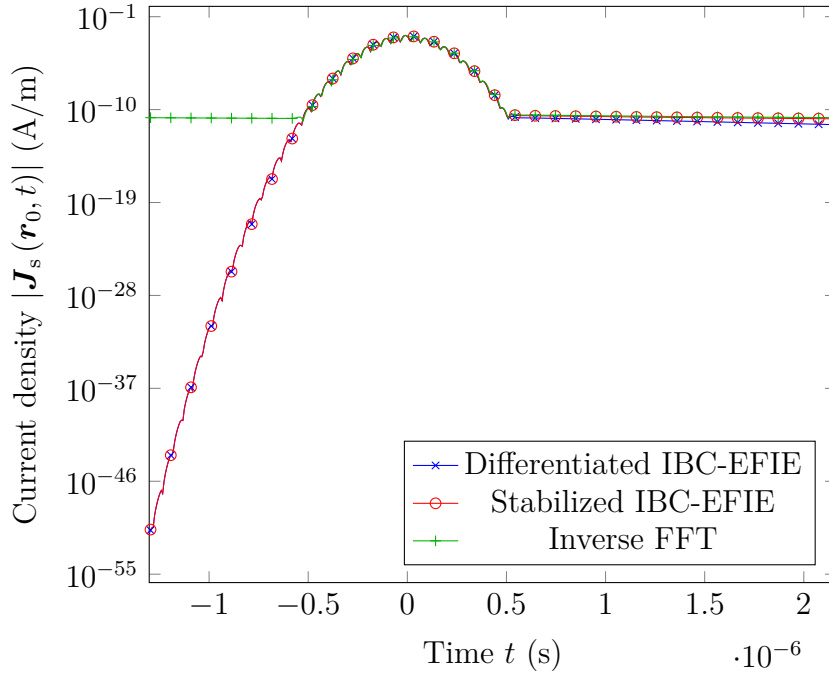


Figure 6.2: Surface current density on the unit sphere

The excitation is a modulated Gaussian plane wave as defined in (5.63)

$$\mathbf{E}^{\text{inc}}(\mathbf{r}, t) = \hat{\mathbf{p}} \exp\left(-\frac{\tau_d(\mathbf{r}, t)^2}{2\tau_{\text{ch}}^2}\right) \cos(2\pi f_0 \tau_d(\mathbf{r}, t)) E_0 \quad (6.60)$$

where $\hat{\mathbf{k}} = -\hat{\mathbf{z}}$, $\tau_d(\mathbf{r}, t) = t - \frac{1}{c} \hat{\mathbf{k}} \cdot \mathbf{r}$, $\hat{\mathbf{p}} = \hat{\mathbf{x}}$, $E_0 = 1$ V/m, $f_0 = 10$ MHz and $\tau_{\text{ch}} = 86.8$ ns. In the figure 6.2, the solution obtained with the differentiated IBC-EFIE (6.26) and the large time step and dense mesh stable IBC-EFIE (6.57) is compared to the reference solution obtained with the inverse FFT of the frequency domain currents. The norm of the current density current is computed in $\mathbf{r}_0 = (0.39 \text{ m}, 0.91 \text{ m}, 0.11 \text{ m})$ for $t \in [-15\tau_{\text{ch}}, 25\tau_{\text{ch}}]$ ($N_t = 512$, $\Delta t = 6.8$ ns). Although, the formulations yield the correct result in the early time time, it seems that they suffers from a DC instability as there is a saturation in the current due to the finite machine precision. The removal of this DC instability will be investigated in the future research.

6.5.2 Stability of the formulation

In this subsection, the condition numbers of the linear systems to solve in the MOTs are studied. The condition number of 3 matrices are computed:

- $\mathbf{S}_0^{\text{der}}$ defined in (6.27a) which is the matrix of the MOT for the differentiated (non-stabilized) IBC-EFIE (6.30),
- $\mathbf{S}_0^{\text{stab}}$ defined in (6.46) which is the matrix of the MOT for the large time step stable IBC-EFIE (6.53),

- $\tilde{\mathbb{T}}_0 \tilde{\mathbf{G}}_m^{-1} \mathbf{S}_0^{\text{stab}}$ which is the matrix of the MOT for the large time step and dense mesh stable IBC-EFIE (6.57).

In the figure 6.3, the condition numbers of the systems and the average number of iterations to solve them (GMRES restarted after every 20 iterations to reach a residual norm of 10^{-8}) have been computed as functions of the time step Δt for a unit sphere ($h \approx 0.2$ m, $N_s = 1080$). As one can expect, the condition number grows proportionally to $\Delta t^{1.5}$ for the non-preconditioned formulation. This is analogous to the growing proportional to $k^{-1.5}$ in the frequency domain when the impedance is proportional to the square root of the frequency and large enough (see (2.30) and figure 3.4). The condition numbers of the stabilized formulations remain bounded as $\Delta t \rightarrow +\infty$. Similarly, the iteration counts keep increasing when the time step increases in the non-stabilized IBC-EFIE while it remains bounded for the stabilized formulations.

In the figure 6.4, the condition numbers and the iteration counts have been computed as functions of the average edge length h on the unit sphere with a fixed time step $\Delta t = 7.16$ ns. The effectiveness of the preconditioner is evident. In fact, with the given set of parameters, the differentiated (non-preconditioned) IBC-EFIE requires more iterations than N_s to converge. Though, it seems that the condition number and the number of iterations for the dense mesh stabilized formulation is not constant as $h \rightarrow 0$ and slowly increases. Future research topics include an analysis of the preconditioned formulation to have a better understanding of this behaviour.

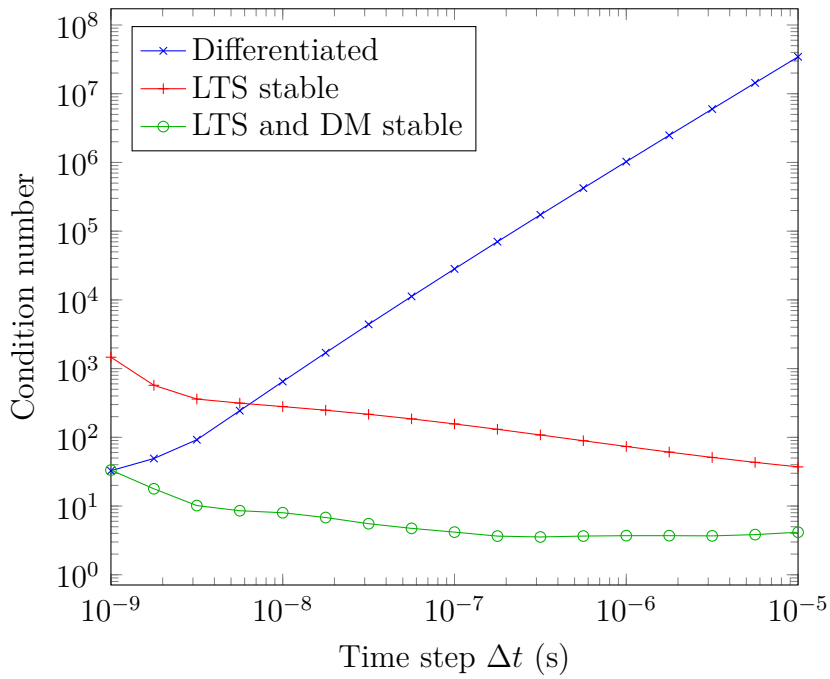
6.6 Conclusion and future work

In this chapter, a new formulation for the time domain IBC-EFIE has been presented. In a first part, it has been shown that the solver must keep track of the losses introduced by the IBC. In a naive formulation this is achieved with an unlimited sequence of interaction matrices, which does not scale well computationally when the number of time steps increases. With the use of an auxiliary current-like variable to compute the RHS of the MOT, the new formulation can take advantage of the lossless operators of the PEC-EFIE and PEC-MFIE that appear in the IBC-EFIE operator, to have a finite number of interaction matrices. This effectively reduces the complexity of the solvers. In a second part, the formulation is stabilized for large time step and for dense meshes. This is achieved with the use of the quasi-Helmholtz projectors and a Calderón-like preconditioning.

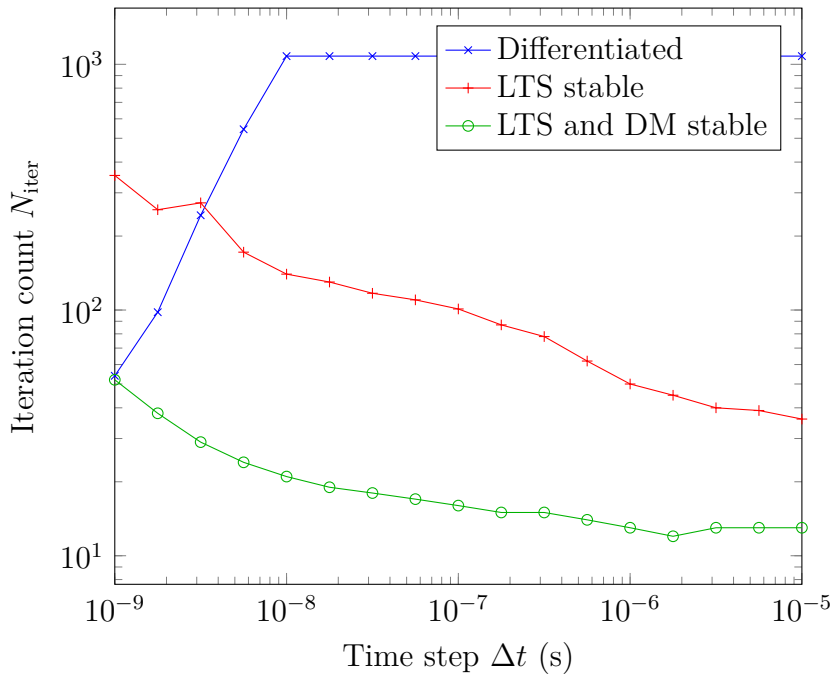
Future research on this time domain IBC-EFIE could include the following topics

- DC instability. As it is now, the formulation still suffers from a numerical instability that seems to be a kind of DC instability as it can be seen in the figure 6.2. Further investigations of this instability may reveal how it can be solved.
- Dense mesh stabilization. Even though the dense mesh preconditioner is effective to significantly reduce the condition number, it seems that it remains

6. Large time step and dense mesh stable time domain IBC-EFIE

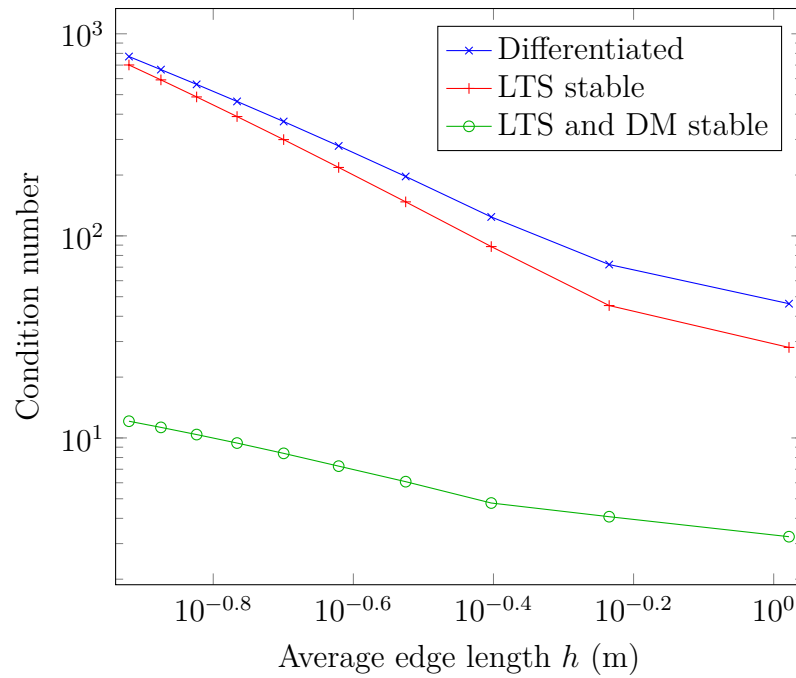


(a) Condition number of the system solved in the MOT

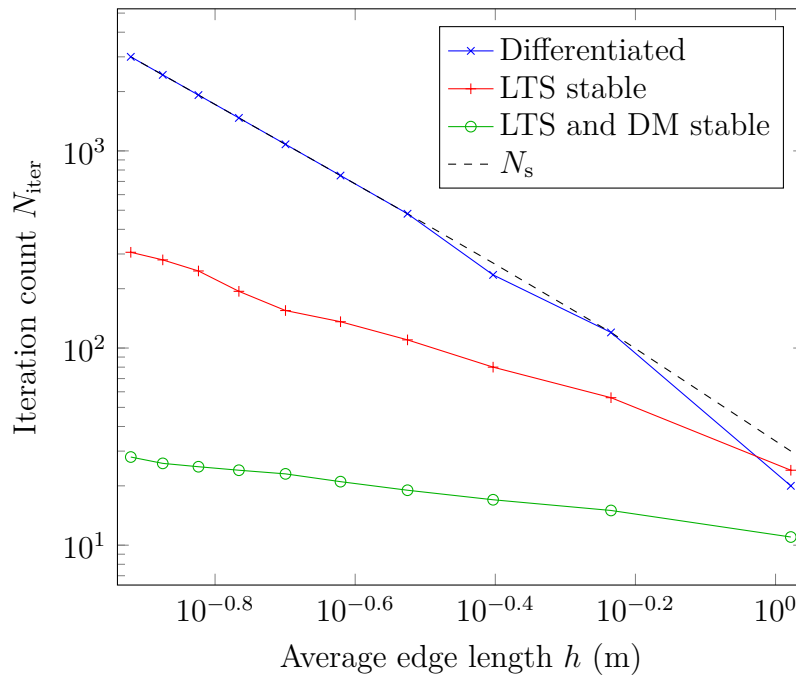


(b) Iteration count required to solve the MOT (limited to $N_s = 1080$ iterations)

Figure 6.3: Condition number and iteration count required to solve the MOT system of the IBC-EFIE as functions of the time step Δt



(a) Condition number of the system solved in the MOT



(b) Iteration count required to solve the MOT

Figure 6.4: Condition number and iteration count required to solve the MOT system of the IBC-EFIE as functions of the edge length h

a slower growth of the condition number when the average edge length h is decreased as it can be observed in the figure 6.4.

- Small time steps (high frequency). As explained in the previous chapter, the compatibility of this technique with the Plane Wave Time Domain (PWTD) algorithm [44, 45] should be investigated. It should be done in addition to the CFIE to remove the resonances.
- Accuracy and numerical stability. Given the large number of terms on the RHS of the MOT (6.54), it should be investigated how to properly reorganize and compute the terms to avoid numerical cancellations and potentially improve the accuracy of the formulation. It might also help to solve the DC instability mentioned before.
- Other formulations in the time domain. Future investigations will focus on using the techniques introduced in this chapter to solve efficiently other problems with losses, and to stabilize them for large time steps and dense mesh.

Conclusions and future work

In this thesis, several computational strategies have been presented to obtain stable and efficient solvers of scattering problems in electromagnetics, both in the frequency domain and in the time domain. The Impedance Boundary Conditions (IBC) and the Electric Field Integral Equation (EFIE) have been the main threads to develop these new techniques.

In the frequency domain part, two low frequency and dense mesh stable IBC-EFIE formulations have been obtained using different approaches. The first Leontovich IBC-EFIE formulation leverages the quasi-Helmholtz projectors to obtain a solver that is stable and free of numerical cancellations. A Calderón-like preconditioning is used to accomplish the dense mesh stabilization. The resulting formulation delivers the correct result for arbitrarily low frequency and dense meshes. The second IBC-EFIE formulation is based on a new IBC that is specifically designed to enable the construction of a multiplicative preconditioner. In the Perfect Electric Conductor (PEC) limit, the formulation remains well-conditioned since it reduces to the classical Calderón preconditioned PEC-EFIE.

In the time domain part, the temporal discretization is achieved using a convolution quadrature based on implicit Runge-Kutta methods. The first time domain formulation is a PEC-EFIE that is stable for large time steps and that is free of the DC instability. The quasi-Helmholtz projectors are used to separate the Helmholtz components of the equation which enables applying the adequate integration or differentiation independently. Finally, the second time domain formulation is an IBC-EFIE. By ensuring that the lossy impedance does not appear in the convolutions of the marching-on-in-time scheme, a computationally efficient algorithm is obtained. Then, a preconditioner based on the one obtained for frequency domain IBC-EFIE is used to stabilize the formulation in the large time step and dense mesh regimes.

Among the potential topics of investigations discussed in the conclusions of each chapter, the future research will be more particularly focused on:

- the investigation on the slow growth of the condition number with the mesh refinement and on the DC instability that is still present in the time domain IBC-EFIE presented in the chapter 6,
- the extension of the preconditioning techniques to non-uniform and anisotropic IBC.

Appendix

Appendix A

Computation of the MoM matrix elements

A.1 Expression of the matrix elements

In order to be computed, the elements of the matrix \mathbf{T}_s (1.104a) are rewritten as

$$[\mathbf{T}_s]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_s \mathbf{f}_n \rangle \quad (\text{A.1a})$$

$$= \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \hat{\mathbf{n}}(\mathbf{r}) \times \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \mathbf{f}_n(\mathbf{r}') dS' dS \quad (\text{A.1b})$$

$$= \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \mathbf{f}_n(\mathbf{r}') dS' dS \quad (\text{A.1c})$$

where the simplification of $\hat{\mathbf{n}}$ between (A.1b) and (A.1c) is valid because \mathbf{f}_i is tangent to Γ .

The elements of \mathbf{T}_h (1.104b) are

$$[\mathbf{T}_h]_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{T}_h \mathbf{f}_n \rangle \quad (\text{A.2a})$$

$$= \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \hat{\mathbf{n}}(\mathbf{r}) \times \nabla \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{f}_n(\mathbf{r}') dS' dS \quad (\text{A.2b})$$

$$= \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \nabla \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{f}_n(\mathbf{r}') dS' dS \quad (\text{A.2c})$$

$$= - \iint_{\mathbf{r} \in \Gamma} \nabla \cdot \mathbf{f}_m(\mathbf{r}) \iint_{\mathbf{r}' \in \Gamma} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{f}_n(\mathbf{r}') dS' dS \quad (\text{A.2d})$$

where the passage from (A.2c) to (A.2d) uses the duality between the gradient and the divergence on Γ .

Finally, the elements of \mathbf{K} are

$$\mathbf{K}_{mn} = \langle \hat{\mathbf{n}} \times \mathbf{f}_m, \mathcal{K} \mathbf{g}_n \rangle \quad (\text{A.3a})$$

$$= \iint_{\mathbf{r} \in \Gamma} \hat{\mathbf{n}}(\mathbf{r}) \times \mathbf{f}_m(\mathbf{r}) \cdot \hat{\mathbf{n}}(\mathbf{r}) \times \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla G(\mathbf{r} - \mathbf{r}') \times \mathbf{g}_n(\mathbf{r}') \, dS' dS \quad (\text{A.3b})$$

$$= \iint_{\mathbf{r} \in \Gamma} \mathbf{f}_m(\mathbf{r}) \cdot \text{p.v.} \iint_{\mathbf{r}' \in \Gamma} \nabla G(\mathbf{r} - \mathbf{r}') \times \mathbf{g}_n(\mathbf{r}') \, dS' dS \quad (\text{A.3c})$$

where the gradient of the Green's function relative to \mathbf{r} is

$$\nabla G(\mathbf{r} - \mathbf{r}') = \nabla \left(\frac{e^{-ik|\mathbf{r} - \mathbf{r}'|}}{4\pi |\mathbf{r} - \mathbf{r}'|} \right) \quad (\text{A.4a})$$

$$= \frac{(1 + ik|\mathbf{r} - \mathbf{r}'|) e^{-ik|\mathbf{r} - \mathbf{r}'|}}{4\pi |\mathbf{r} - \mathbf{r}'|^3} (\mathbf{r}' - \mathbf{r}). \quad (\text{A.4b})$$

As it has been shown in the section 1.3.3, the BC basis functions are a linear combination of RWG basis functions in the barycentric refined mesh. Also, an RWG basis function and its divergence are linear combinations of two functions in the form (see (1.90))

$$\mathbf{f}(\mathbf{r}) = \frac{\mathbf{r} - \mathbf{v}}{2A} \quad (\text{A.5a})$$

$$\nabla \cdot \mathbf{f}(\mathbf{r}) = \frac{1}{A} \quad (\text{A.5b})$$

on each triangle of their support. So, each element of the above matrices in the form $\langle \mathbf{t}, \mathcal{L} \mathbf{s} \rangle$ where \mathbf{t} and \mathbf{s} are the testing and source basis functions and \mathcal{L} is the operator, can be computed with the appropriate linear combination of inner products in the form $\langle \mathbf{f}, \mathcal{L} \mathbf{f}' \rangle$ (or $\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{L} \mathbf{f}' \rangle$) for each pair of triangle (T, T') . Here, T is in the support of \mathbf{t} , T' is in the support of \mathbf{s} , and \mathbf{f} and \mathbf{f}' are functions in the form (A.5a) defined respectively on T and T' . In the following, the vertex and the triangle area in (A.5a) are primed for \mathbf{f}' , so they are noted \mathbf{v}' and A' .

A.2 Integration rules

There are several ways to compute the integral \mathbf{I} of a function \mathbf{x} on a triangle T

$$\mathbf{I} = \iint_{\mathbf{r} \in T} \mathbf{x}(\mathbf{r}) \, dS. \quad (\text{A.6})$$

The first way is to compute it analytically. This is usually possible for simple integrals such as the one involved in the computation of the Gram or mix-Gram matrices.

The second way is to use a Gaussian quadrature on the triangle T [78]. A Gaussian quadrature rule $(w_i, \mathbf{r}_i)_{i=1}^N$ is defined by N points $(\mathbf{r}_i)_{i=1}^N$ in the integration domain T and N weights $(w_i)_{i=1}^N$. Then, the integral is approximated as

$$\mathbf{I} \approx \sum_{i=1}^N \mathbf{x}(\mathbf{r}_i) w_i. \quad (\text{A.7})$$

The equality is strict for all polynomials \mathbf{x} up to a certain degree that depends on the order of the rule. When \mathbf{x} is not a polynomial, the Gaussian quadrature usually works well when the integrand \mathbf{x} is not singular. This is the case for far interacting testing and source basis functions. For example the element of \mathbf{T}_s and \mathbb{T}_s can be approximated by a combination of inner products in the form $\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{T} \mathbf{f}' \rangle$ as explained above, then using two Gaussian rules $(w_i, \mathbf{r}_i)_{i=1}^N$ on T and $(w'_j, \mathbf{r}'_j)_{j=1}^{N'}$ on T' ,

$$\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{T}_s \mathbf{f}' \rangle = \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \iint_{\mathbf{r}' \in T'} G(\mathbf{r} - \mathbf{r}') \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.8a})$$

$$\approx \sum_{i=1}^N \mathbf{f}(\mathbf{r}_i) \cdot \sum_{j=1}^{N'} G(\mathbf{r}_i - \mathbf{r}'_j) \mathbf{f}'(\mathbf{r}'_j) w'_j w_i. \quad (\text{A.8b})$$

Similarly, for \mathbf{T}_h and \mathbb{T}_h ,

$$\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{T}_h \mathbf{f}' \rangle = - \iint_{\mathbf{r} \in T} \nabla \cdot \mathbf{f}(\mathbf{r}) \iint_{\mathbf{r}' \in T'} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.9a})$$

$$= - \frac{1}{AA'} \iint_{\mathbf{r} \in T} \iint_{\mathbf{r}' \in T'} G(\mathbf{r} - \mathbf{r}') dS' dS \quad (\text{A.9b})$$

$$\approx - \frac{1}{AA'} \sum_{i=1}^N \sum_{j=1}^{N'} G(\mathbf{r}_i - \mathbf{r}'_j) w'_j w_i. \quad (\text{A.9c})$$

For \mathbf{K} ,

$$\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{K} \mathbf{f}' \rangle = \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \iint_{\mathbf{r}' \in T'} \nabla G(\mathbf{r} - \mathbf{r}') \times \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.10a})$$

$$\approx \sum_{i=1}^N \mathbf{f}(\mathbf{r}_i) \cdot \sum_{j=1}^{N'} \nabla G(\mathbf{r}_i - \mathbf{r}'_j) \times \mathbf{f}'(\mathbf{r}'_j) w'_j w_i. \quad (\text{A.10b})$$

For \mathbf{G} ,

$$\text{if } T \neq T', \langle \mathbf{f}, \mathbf{f}' \rangle = 0 \quad (\text{A.11a})$$

$$\text{if } T = T', \langle \mathbf{f}, \mathbf{f}' \rangle = \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \mathbf{f}'(\mathbf{r}) dS \quad (\text{A.11b})$$

$$\approx \sum_{i=1}^N \mathbf{f}(\mathbf{r}_i) \cdot \mathbf{f}'(\mathbf{r}_i) w_i. \quad (\text{A.11c})$$

For \mathbf{G}_m ,

$$\text{if } T \neq T', \langle \hat{\mathbf{n}} \times \mathbf{f}, \mathbf{f}' \rangle = 0 \quad (\text{A.12a})$$

$$\text{if } T = T', \langle \hat{\mathbf{n}} \times \mathbf{f}, \mathbf{f}' \rangle = \iint_{\mathbf{r} \in T} (\hat{\mathbf{n}} \times \mathbf{f}(\mathbf{r})) \cdot \mathbf{f}'(\mathbf{r}) dS \quad (\text{A.12b})$$

$$\approx \sum_{i=1}^N (\hat{\mathbf{n}} \times \mathbf{f}(\mathbf{r}_i)) \cdot \mathbf{f}'(\mathbf{r}_i) w_i. \quad (\text{A.12c})$$

Finally, when the supports T and T' are close to each other or coincide, the elements of \mathbf{T} , \mathbb{T} and \mathbf{K} cannot be computed directly with Gaussian quadrature rules because the integration kernel is singular.

A.3 Singularity extraction

Several techniques exist to deal with these singular kernels [27, 79–82]. The singularity extraction technique described in the following is standard. It consists in rewriting the integral kernels as a regular part and a singular part. The regular part is integrated using the double Gaussian rules, and the singular part is integrated analytically for the inner integral and uses the Gaussian quadrature for the outer integral. The kernel for \mathbf{T} is

$$G(\mathbf{r} - \mathbf{r}') = \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r} - \mathbf{r}'|} = \underbrace{\frac{e^{-ik|\mathbf{r}-\mathbf{r}'|} - 1}{4\pi|\mathbf{r} - \mathbf{r}'|}}_{G_{\text{reg}}(\mathbf{r} - \mathbf{r}')} + \frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|}. \quad (\text{A.13})$$

A Taylor expansion of the first term $G_{\text{reg}}(\mathbf{r} - \mathbf{r}')$ confirms that it is not singular

$$G_{\text{reg}}(\mathbf{r} - \mathbf{r}') = G(\mathbf{r} - \mathbf{r}') - \frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|} \quad (\text{A.14a})$$

$$= \frac{1 - ik|\mathbf{r} - \mathbf{r}'| + O(|\mathbf{r} - \mathbf{r}'|^2)}{4\pi|\mathbf{r} - \mathbf{r}'|} - \frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|} \quad (\text{A.14b})$$

$$= -\frac{ik}{4\pi} + O(|\mathbf{r} - \mathbf{r}'|). \quad (\text{A.14c})$$

Then, the following integrals are computed analytically for any $\mathbf{r} \in \mathbb{R}^3$ using the method described in [83, 84]

$$I_{1/R}(\mathbf{r}) = \iint_{\mathbf{r}' \in T'} \frac{1}{|\mathbf{r} - \mathbf{r}'|} dS' \quad (\text{A.15a})$$

$$\mathbf{I}_{\mathbf{r}'/R}(\mathbf{r}) = \iint_{\mathbf{r}' \in T'} \frac{\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} dS'. \quad (\text{A.15b})$$

Using these notations, the singular part of the elements for \mathbf{T}_s is

$$\iint_{\mathbf{r}' \in T'} \frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|} \mathbf{f}'(\mathbf{r}') dS' = \iint_{\mathbf{r}' \in T'} \frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|} \frac{\mathbf{r}' - \mathbf{v}'}{2A'} dS' \quad (\text{A.16a})$$

$$= \frac{1}{8\pi A'} (\mathbf{I}_{\mathbf{r}'/R}(\mathbf{r}) - I_{1/R}(\mathbf{r}) \mathbf{v}') \quad (\text{A.16b})$$

and the singular part of the elements for \mathbf{T}_h is

$$\iint_{\mathbf{r}' \in T'} \frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|} \nabla' \cdot \mathbf{f}'(\mathbf{r}') dS' = \frac{1}{4\pi A'} I_{1/R}(\mathbf{r}). \quad (\text{A.17})$$

So (A.8a) becomes

$$\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{T}_s \mathbf{f}' \rangle = \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \iint_{\mathbf{r}' \in T'} G(\mathbf{r} - \mathbf{r}') \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.18a})$$

$$= \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \iint_{\mathbf{r}' \in T'} G_{\text{reg}}(\mathbf{r} - \mathbf{r}') \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.18b})$$

$$+ \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \frac{1}{8\pi A'} (\mathbf{I}_{\mathbf{r}'/R}(\mathbf{r}) - I_{1/R}(\mathbf{r}) \mathbf{v}') dS$$

$$\approx \sum_{i=1}^N \mathbf{f}(\mathbf{r}_i) \cdot \sum_{j=1}^{N'} G_{\text{reg}}(\mathbf{r}_i - \mathbf{r}'_j) \mathbf{f}'(\mathbf{r}'_j) w'_j w_i \quad (\text{A.18c})$$

$$+ \frac{1}{8\pi A'} \sum_{i=1}^N \mathbf{f}(\mathbf{r}_i) \cdot (\mathbf{I}_{\mathbf{r}'/R}(\mathbf{r}_i) - I_{1/R}(\mathbf{r}_i) \mathbf{v}') w_i$$

and (A.9a) becomes

$$\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{T}_h \mathbf{f}' \rangle = - \iint_{\mathbf{r} \in T} \nabla \cdot \mathbf{f}(\mathbf{r}) \iint_{\mathbf{r}' \in T'} G(\mathbf{r} - \mathbf{r}') \nabla' \cdot \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.19a})$$

$$= - \frac{1}{AA'} \iint_{\mathbf{r} \in T} \iint_{\mathbf{r}' \in T'} G_{\text{reg}}(\mathbf{r} - \mathbf{r}') dS' dS \quad (\text{A.19b})$$

$$- \frac{1}{4\pi AA'} \iint_{\mathbf{r} \in T} I_{1/R}(\mathbf{r}) dS$$

$$\approx - \frac{1}{AA'} \sum_{i=1}^N \sum_{j=1}^{N'} G_{\text{reg}}(\mathbf{r}_i - \mathbf{r}'_j) w'_j w_i \quad (\text{A.19c})$$

$$- \frac{1}{4\pi AA'} \sum_{i=1}^N I_{1/R}(\mathbf{r}_i) w_i.$$

Similarly, the kernel for \mathbf{K} is the gradient of the Green's function (A.4). In the following $R = |\mathbf{r} - \mathbf{r}'|$,

$$\nabla G(\mathbf{r} - \mathbf{r}') = \frac{(1 + ikR) e^{-ikR}}{4\pi R^3} (\mathbf{r}' - \mathbf{r}) \quad (\text{A.20a})$$

$$= \underbrace{\frac{(1 + ikR) e^{-ikR} - 1 - \frac{1}{2} k^2 R^2}{4\pi R^3}}_{(\nabla G)_{\text{reg}}(\mathbf{r} - \mathbf{r}')} (\mathbf{r}' - \mathbf{r}) + \left(\frac{1}{4\pi R^3} + \frac{k^2}{8\pi R} \right) (\mathbf{r}' - \mathbf{r}).$$

$$(\text{A.20b})$$

A. Computation of the MoM matrix elements

Again, a Taylor expansion of the regular part of the gradient of the Green's function $(\nabla G)_{\text{reg}}(\mathbf{r} - \mathbf{r}')$ shows that it has no singular contributions

$$(\nabla G)_{\text{reg}}(\mathbf{r} - \mathbf{r}') = \nabla G(\mathbf{r} - \mathbf{r}') - \left(\frac{1}{4\pi R^3} + \frac{k^2}{8\pi R} \right) (\mathbf{r}' - \mathbf{r}) \quad (\text{A.21a})$$

$$= \left(\frac{1 + \frac{1}{2}k^2 R^2 - \frac{1}{3}ik^3 R^3 + O(R^4)}{4\pi R^3} - \frac{1}{4\pi R^3} - \frac{k^2}{8\pi R} \right) (\mathbf{r}' - \mathbf{r}) \quad (\text{A.21b})$$

$$= \left(-\frac{ik^3}{12\pi} + O(R) \right) (\mathbf{r}' - \mathbf{r}). \quad (\text{A.21c})$$

Similarly to (A.15), the following integrals are computed analytically for any $\mathbf{r} \in \mathbb{R}^3$ using the method described in [83, 84]

$$I_{1/R^3}(\mathbf{r}) = \iint_{\mathbf{r}' \in T'} \frac{1}{|\mathbf{r} - \mathbf{r}'|^3} dS' \quad (\text{A.22a})$$

$$\mathbf{I}_{\mathbf{r}'/R^3}(\mathbf{r}) = \iint_{\mathbf{r}' \in T'} \frac{\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dS'. \quad (\text{A.22b})$$

The singular part of the elements for \mathbf{K} that is computed analytically is

$$\mathbf{K}_{\text{an}}(\mathbf{r}) = \iint_{\mathbf{r}' \in T'} \left(\frac{1}{4\pi R^3} + \frac{k^2}{8\pi R} \right) (\mathbf{r}' - \mathbf{r}) \times \mathbf{f}'(\mathbf{r}') dS' \quad (\text{A.23a})$$

$$= \frac{1}{8\pi A'} \iint_{\mathbf{r}' \in T'} \left(\frac{1}{R^3} + \frac{k^2}{2R} \right) (\mathbf{r}' - \mathbf{r}) \times (\mathbf{r}' - \mathbf{v}') dS' \quad (\text{A.23b})$$

$$= \frac{1}{8\pi A'} \iint_{\mathbf{r}' \in T'} \left(\frac{1}{R^3} + \frac{k^2}{2R} \right) (\mathbf{r} \times \mathbf{v}' + (\mathbf{v}' - \mathbf{r}) \times \mathbf{r}') dS' \quad (\text{A.23c})$$

$$= \frac{1}{8\pi A'} \left(I_{1/R^3}(\mathbf{r}) + \frac{k^2}{2} I_{1/R}(\mathbf{r}) \right) \mathbf{r} \times \mathbf{v}' \quad (\text{A.23d})$$

$$+ \frac{1}{8\pi A'} (\mathbf{v}' - \mathbf{r}) \times \left(\mathbf{I}_{\mathbf{r}'/R^3}(\mathbf{r}) + \frac{k^2}{2} \mathbf{I}_{\mathbf{r}'/R}(\mathbf{r}) \right).$$

So (A.10a) becomes

$$\langle \hat{\mathbf{n}} \times \mathbf{f}, \mathcal{K} \mathbf{f}' \rangle = \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \iint_{\mathbf{r}' \in T'} \nabla G(\mathbf{r} - \mathbf{r}') \times \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.24a})$$

$$= \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \iint_{\mathbf{r}' \in T'} (\nabla G)_{\text{reg}}(\mathbf{r} - \mathbf{r}') \times \mathbf{f}'(\mathbf{r}') dS' dS \quad (\text{A.24b})$$

$$+ \iint_{\mathbf{r} \in T} \mathbf{f}(\mathbf{r}) \cdot \mathbf{K}_{\text{an}}(\mathbf{r}) dS$$

$$\approx \sum_{i=1}^N \mathbf{f}(\mathbf{r}_i) \cdot \sum_{j=1}^{N'} (\nabla G)_{\text{reg}}(\mathbf{r}_i - \mathbf{r}'_j) \times \mathbf{f}'(\mathbf{r}'_j) w'_j w_i \quad (\text{A.24c})$$

$$+ \sum_{i=1}^N \mathbf{f}(\mathbf{r}_i) \cdot \mathbf{K}_{\text{an}}(\mathbf{r}_i) w_i.$$

Acronyms and notations

Acronym	Meaning
BC	Buffa-Christiansen
BEM	Boundary Element Method
CFIE	Combined Field Integral Equation
CN	Condition Number
DC	Direct Current
DM	Dense Mesh
EFIE	Electric Field Integral Equation
e.g.	for example
FD	Frequency Domain
FFT	Fast Fourier Transform
FMM	Fast Multipole Method
GMRES	General Minimal Residual
IBC	Impedance Boundary Condition
i.e.	in other words
IRK	Implicit Runge-Kutta
LF	Low Frequency
LHS	Left Hand Side
LTS	Large Time Step
MFIE	Magnetic Field Integral Equation
MLFMM	Multi-Level Fast Multipole Method
MoM	Method of Moments
MOT	Marching-On-in-Time
PEC	Perfect Electric Conductor
PW	Plane Wave
RCS	Radar Cross Section
RHS	Right Hand Side
RWG	Rao-Wilton-Glisson
SVD	Singular Value Decomposition
TD	Time Domain
VSH	Vector Spherical Harmonics

Table A.1: Common acronyms

Notations	Description
r, s	Scalar
\mathbf{r}, \mathbf{E}	Geometric vector in \mathbb{R}^3 or \mathbb{C}^3
$\hat{\mathbf{r}}, \hat{\mathbf{u}}$	Unit vector in \mathbb{R}^3
\mathbf{A}, \mathbf{v}	1-dimensional array, column vector in \mathbb{R}^n or \mathbb{C}^n
\mathbf{A}, \mathbf{s}	2-dimensional array, matrix in $\mathbb{R}^{m \times n}$ or $\mathbb{C}^{m \times n}$
\mathcal{L}, \mathcal{T}	Linear operator
e	Euler's number ($e \approx 2.718$)
i	Imaginary unit ($i^2 = -1$)
$\Re(z), \Im(z)$	Real part, imaginary part
$ z , \mathbf{r} $	Absolute value, Norm 2 (Euclidean) for geometric vectors
$\ \mathbf{M}\ , \ \mathbf{E}\ $	Norm 2 for matrices and column vectors
$\mathbf{v}^\top, \mathbf{M}^\top$	Transpose
$\overline{\mathbf{M}}, \overline{\mathbf{f}}$	Conjugate
\mathcal{T}^*	Adjoint operator
$v_n, [\mathbf{M}\mathbf{v}]_n$	n -th element of a vector
x_i	i -th element of a sequence
$\mathbf{M}_{mn}, \mathbf{M}_{m,n}, [\mathbf{M}\mathbf{N}]_{mn}$	Element (m, n) of a matrix
$\langle \mathbf{f}, \mathbf{g} \rangle$	Inner product
$\mathbf{f} \times \mathbf{g}$	Cross product
$\mathbf{f} \cdot \mathbf{g}$	Dot product
$\mathbf{M} \otimes \mathbf{N}$	Kronecker product
$(f * g)(t), (x * y)_n$	Temporal convolution, Sequence convolution
$\frac{d}{dx}, f'(x)$	Derivative
$\frac{\partial}{\partial t}$	Partial derivative
$\nabla f, \nabla' f$	Gradient (relative to \mathbf{r}), Gradient relative to \mathbf{r}'
$\nabla \cdot \mathbf{f}, \nabla' \cdot \mathbf{f}$	Divergence (relative to \mathbf{r}), Divergence relative to \mathbf{r}'
$\nabla \times \mathbf{f}$	Curl
$\nabla^2 f, \nabla^2 \mathbf{f}$	Laplacian, Vector Laplacian
$\partial V, \partial \Gamma$	Boundary of a domain

Table A.2: Common notations

Publications

- [39] A. Dély, F. P. Andriulli, and K. Cools, “An impedance boundary condition EFIE that is low-frequency and refinement stable,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 3, pp. 1259–1266, 2017.
- [40] A. Dély, F. P. Andriulli, and K. Cools, “On a well-conditioned impedance boundary condition EFIE,” in *IEEE International Symposium on Antennas and Propagation*, 2017, pp. 715–716.
- [42] A. Dély, F. P. Andriulli, and K. Cools, “On a conforming impedance boundary condition EFIE,” in *IEEE International Symposium on Antennas and Propagation*, 2018.
- [70] A. Dély, F. P. Andriulli, and K. Cools, “Stable TD-EFIE discretized with implicit Runge-Kutta methods,” in *European Conference on Antennas and Propagation*, 2018.

Bibliography

- [1] J.-M. Jin, *Theory and computation of electromagnetic fields*. John Wiley & Sons, 2011.
- [2] J.-C. Nédélec, *Acoustic and electromagnetic equations: integral representations for harmonic problems*. Springer Science, 2001, vol. 144.
- [3] S. Rao, D Wilton, and A. Glisson, “Electromagnetic scattering by surfaces of arbitrary shape,” *IEEE Transactions on antennas and propagation*, vol. 30, no. 3, pp. 409–418, 1982.
- [4] A. Buffa and S. Christiansen, “A dual finite element complex on the barycentric refinement,” *Mathematics of Computation*, vol. 76, no. 260, pp. 1743–1769, 2007.
- [5] G. Vecchi, “Loop-star decomposition of basis functions in the discretization of the EFIE,” *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 2, pp. 339–346, 1999.
- [6] K. Cools, F. Andriulli, D. De Zutter, and E Michielssen, “Accurate and conforming mixed discretization of the MFIE,” *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 528–531, 2011.
- [7] F. P. Andriulli, K. Cools, H. Bagci, F. Olyslager, A. Buffa, S. Christiansen, and E. Michielssen, “A multiplicative Calderon preconditioner for the electric field integral equation,” *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 8, pp. 2398–2412, 2008.
- [8] A. J. Poggio and E. K. Miller, “Integral equation solutions of three-dimensional scattering problems,” in *Computer techniques for electromagnetics*, Pergamon Press, 1973, ch. 4.
- [9] Y. Chang and R Harrington, “A surface formulation for characteristic modes of material bodies,” *IEEE Transactions on Antennas and Propagation*, vol. 25, no. 6, pp. 789–795, 1977.
- [10] T.-K. Wu and L. L. Tsai, “Scattering from arbitrarily-shaped lossy dielectric bodies of revolution,” *Radio Science*, vol. 12, no. 5, pp. 709–718, 1977.
- [11] C. Müller, *Foundations of the mathematical theory of electromagnetic waves*. Springer, 2013, vol. 155.
- [12] D. J. Hoppe, *Impedance boundary conditions in electromagnetics*. CRC Press, 1995.

-
- [13] M. Leontovich, "Approximate boundary conditions for the electromagnetic field on the surface of a good conductor," *Investigations on Radiowave Propagation*, vol. 2, pp. 5–12, 1948.
- [14] T. Senior, "Impedance boundary conditions for imperfectly conducting surfaces," *Applied Scientific Research, Section B*, vol. 8, no. 1, p. 418, 1960.
- [15] N. G. Alexopoulos and G. A. Tadler, "Accuracy of the Leontovich boundary condition for continuous and discontinuous surface impedances," *Journal of Applied Physics*, vol. 46, no. 8, pp. 3326–3332, 1975.
- [16] D.-S. Wang, "Limits and validity of the impedance boundary condition on penetrable surfaces," *IEEE Transactions on Antennas and Propagation*, vol. 35, no. 4, pp. 453–457, 1987.
- [17] S. Yuferev and N. Ida, "Selection of the surface impedance boundary conditions for a given problem," *IEEE Transactions on Magnetics*, vol. 35, no. 3, pp. 1486–1489, 1999.
- [18] J.-R. Poirier, A. Bendali, and P. Borderies, "Impedance boundary conditions for the scattering of time-harmonic waves by rapidly varying surfaces," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 3, pp. 995–1005, 2006.
- [19] M. L. Goodman, "Scattering by a sphere with an anisotropic impedance boundary condition: Mie series solution and uniqueness conditions," *Journal of the Optical Society of America A*, vol. 7, no. 10, pp. 1817–1823, 1990.
- [20] A. W. Glisson, M. Orman, F. Falco, and D. Koppel, "Electromagnetic scattering by an arbitrarily shaped surface with an anisotropic impedance boundary condition," *Applied Computational Electromagnetics Society Journal*, vol. 10, no. 3, pp. 93–106, 1995.
- [21] I. V. Lindell and A. H. Sihvola, "Realization of impedance boundary," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 12, pp. 3669–3676, 2006.
- [22] D. J. Hoppe and Y. Rahmat-Samii, "Scattering by superquadric dielectric-coated cylinders using higher order impedance boundary conditions," *IEEE Transactions on Antennas and Propagation*, vol. 40, no. 12, pp. 1513–1523, 1992.
- [23] R. Cicchetti, "A class of exact and higher-order surface boundary conditions for layered structures," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 2, pp. 249–259, 1996.
- [24] O. Marceaux and B. Stupfel, "High-order impedance boundary conditions for multilayer coated 3-D objects," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 3, pp. 429–436, 2000.
- [25] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas and Propagation Magazine*, vol. 35, no. 3, pp. 7–12, 1993.

- [26] J. Song and W. C. Chew, “Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering,” *Microwave and Optical Technology Letters*, vol. 10, no. 1, pp. 14–19, 1995.
- [27] W. C. Chew, E. Michielssen, J. Song, and J.-M. Jin, *Fast and efficient algorithms in computational electromagnetics*. Artech House, Inc., 2001.
- [28] G. C. Hsiao and R. E. Kleinman, “Mathematical foundations for error estimation in numerical solutions of integral equations in electromagnetics,” *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 3, pp. 316–328, 1997.
- [29] H. Contopanagos, B. Dembart, M. Epton, J. J. Ottusch, V. Rokhlin, J. L. Visher, and S. M. Wandzura, “Well-conditioned boundary integral equations for three-dimensional electromagnetic scattering,” *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 12, pp. 1824–1830, 2002.
- [30] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST handbook of Mathematical Functions*. Cambridge University Press, 2010.
- [31] F. P. Andriulli, “Loop-star and loop-tree decompositions: Analysis and efficient algorithms,” *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 5, pp. 2347–2356, 2012.
- [32] A. Bendali, M. Fares, and J. Gay, “A boundary-element solution of the Leontovitch problem,” *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 10, pp. 1597–1605, 1999.
- [33] Z. G. Qian, W. C. Chew, and R. Suaya, “Generalized impedance boundary condition for conductor modeling in surface integral equation,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 11, pp. 2354–2364, 2007.
- [34] P. Ylä-Oijala, S. P. Kiminki, and S. Jarvenpaa, “Solving IBC-CFIE with dual basis functions,” *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 12, pp. 3997–4004, 2010.
- [35] S. Yan and J.-M. Jin, “Self-dual integral equations for electromagnetic scattering from IBC objects,” *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 11, pp. 5533–5546, 2013.
- [36] P. Ylä-Oijala, S. P. Kiminki, H. Wallén, and A. Sihvola, “Uniform surface integral equation formulation for mixed impedance boundary conditions,” *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 12, pp. 5718–5726, 2015.
- [37] A. W. Glisson, “Electromagnetic scattering by arbitrarily shaped surfaces with impedance boundary conditions,” *Radio Science*, vol. 27, no. 06, pp. 935–943, 1992.
- [38] F. P. Andriulli, K. Cools, I. Bogaert, and E. Michielssen, “On a well-conditioned electric field integral operator for multiply connected geometries,” *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 2077–2087, 2013.

-
- [39] A. Dély, F. P. Andriulli, and K. Cools, “An impedance boundary condition EFIE that is low-frequency and refinement stable,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 3, pp. 1259–1266, 2017.
- [40] A. Dély, F. P. Andriulli, and K. Cools, “On a well-conditioned impedance boundary condition EFIE,” in *IEEE International Symposium on Antennas and Propagation*, 2017, pp. 715–716.
- [41] J. D. Jackson, “Classical electrodynamics,” 1998.
- [42] A. Dély, F. P. Andriulli, and K. Cools, “On a conforming impedance boundary condition EFIE,” in *IEEE International Symposium on Antennas and Propagation*, 2018.
- [43] A. E. Yilmaz, J.-M. Jin, and E. Michielssen, “Time domain adaptive integral method for surface integral equations,” *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 10, pp. 2692–2708, 2004.
- [44] A. A. Ergin, B. Shanker, and E. Michielssen, “The plane-wave time-domain algorithm for the fast analysis of transient wave phenomena,” *IEEE Antennas and Propagation Magazine*, vol. 41, no. 4, pp. 39–52, 1999.
- [45] B. Shanker, A. A. Ergin, M. Lu, and E. Michielssen, “Fast analysis of transient electromagnetic scattering phenomena using the multilevel plane wave time domain algorithm,” *IEEE Transactions on Antennas and Propagation*, vol. 51, no. 3, pp. 628–641, 2003.
- [46] J. C. Butcher, *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.
- [47] L. Banjai, M. Messner, and M. Schanz, “Runge–Kutta convolution quadrature for the boundary element method,” *Computer Methods in Applied Mechanics and Engineering*, vol. 245, pp. 90–101, 2012.
- [48] C. Lubich and A. Ostermann, “Runge–Kutta methods for parabolic equations and convolution quadrature,” *Mathematics of Computation*, vol. 60, no. 201, pp. 105–131, 1993.
- [49] F.-J. Sayas, *Retarded Potentials and Time Domain Boundary Integral Equations: A Road Map*. Springer, 2016, vol. 50.
- [50] F. P. Andriulli, K. Cools, F. Olyslager, and E. Michielssen, “Time domain Calderón identities and their application to the integral equation analysis of scattering by PEC objects Part II: Stability,” *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 8, pp. 2365–2375, 2009.
- [51] E. Miller, “Time-domain modeling in electromagnetics,” *Journal of Electromagnetic Waves and Applications*, vol. 8, no. 9-10, pp. 1125–1172, Jan. 1994.
- [52] A. E. Yilmaz, D. S. Weile, H.-M. Jin, and E. Michielssen, “A hierarchical FFT algorithm (HIL-FFT) for the fast analysis of transient electromagnetic scattering phenomena,” *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 7, pp. 971–982, 2002.

- [53] N.-W. Chen, K Aygün, and E Michielssen, “Integral-equation-based analysis of transient scattering and radiation from conducting bodies at very low frequencies,” *IEE Proceedings-Microwaves, Antennas and Propagation*, vol. 148, no. 6, pp. 381–387, 2001.
- [54] G. Pisharody and D. S. Weile, “Robust solution of time-domain integral equations using loop-tree decomposition and bandlimited extrapolation,” *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 6, pp. 2089–2098, 2005.
- [55] F. P. Andriulli, H. Bagci, F. Vipiana, G. Vecchi, and E. Michielssen, “Analysis and regularization of the TD-EFIE low-frequency breakdown,” *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 7, pp. 2034–2046, 2009.
- [56] Y. Beghein, K. Cools, and F. P. Andriulli, “A DC stable and large-time step well-balanced TD-EFIE based on quasi-Helmholtz projectors,” *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 7, pp. 3087–3097, 2015.
- [57] B. Shanker, A. A. Ergin, K. Aygun, and E. Michielssen, “Analysis of transient electromagnetic scattering from closed surfaces using a combined field integral equation,” *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 7, pp. 1064–1074, 2000.
- [58] Y. Beghein, K. Cools, H. Bagci, and D. De Zutter, “A space-time mixed Galerkin marching-on-in-time scheme for the time-domain combined field integral equation,” *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 3, pp. 1228–1238, 2013.
- [59] R. A. Wildman, G. Pisharody, D. S. Weile, S. Balasubramaniam, and E. Michielssen, “An accurate scheme for the solution of the time-domain integral equations of electromagnetics using higher order vector bases and bandlimited extrapolation,” *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 11, pp. 2973–2984, 2004.
- [60] R. D. Graglia, D. R. Wilton, and A. F. Peterson, “Higher order interpolatory vector bases for computational electromagnetics,” *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 3, pp. 329–342, 1997.
- [61] J.-L. Hu, C. H. Chan, and Y. Xu, “A new temporal basis function for the time-domain integral equation method,” *IEEE Microwave and Wireless Components Letters*, vol. 11, no. 11, pp. 465–466, 2001.
- [62] D. S. Weile, G. Pisharody, N.-W. Chen, B. Shanker, and E. Michielssen, “A novel scheme for the solution of the time-domain integral equations of electromagnetics,” *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 1, pp. 283–295, 2004.
- [63] E. van’t Wout, D. R. van der Heul, H. van der Ven, and C. Vuik, “Design of temporal basis functions for time domain integral equation methods with predefined accuracy and smoothness,” *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 1, pp. 271–280, 2013.

-
- [64] B. Shanker, M. Lu, J. Yuan, and E. Michielssen, “Time domain integral equation analysis of scattering from composite bodies via exact evaluation of radiation fields,” *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 5, pp. 1506–1520, 2009.
- [65] Y. Shi, M.-Y. Xia, R.-S. Chen, E. Michielssen, and M. Lu, “Stable electric field TDIE solvers via quasi-exact evaluation of MOT matrix elements,” *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 2, pp. 574–585, 2011.
- [66] Y.-S. Chung, T. K. Sarkar, B. H. Jung, M. Salazar-Palma, Z. Ji, S. Jang, and K. Kim, “Solution of time domain electric field integral equation using the Laguerre polynomials,” *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 9, pp. 2319–2328, 2004.
- [67] A. J. Pray, Y. Beghein, N. V. Nair, K. Cools, H. Bağcı, and B. Shanker, “A higher order space-time Galerkin scheme for time domain integral equations,” *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 12, pp. 6183–6191, 2014.
- [68] X. Wang and D. S. Weile, “Implicit Runge-Kutta methods for the discretization of time domain integral equations,” *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 12, pp. 4651–4663, 2011.
- [69] Q. Chen, P. Monk, X. Wang, and D. Weile, “Analysis of convolution quadrature applied to the time-domain electric field integral equation,” *Communications in Computational Physics*, vol. 11, no. 2, pp. 383–399, 2012.
- [70] A. Dély, F. P. Andriulli, and K. Cools, “Stable TD-EFIE discretized with implicit Runge-Kutta methods,” in *European Conference on Antennas and Propagation*, 2018.
- [71] S. Yuferev and N. Ida, “Time domain surface impedance concept for low frequency electromagnetic problems-part I: Derivation of high order surface impedance boundary conditions in the time domain,” *IEE Proceedings-Science, Measurement and Technology*, vol. 152, no. 4, pp. 175–185, 2005.
- [72] F. M. Tesche, “On the inclusion of loss in time-domain solutions of electromagnetic interaction problems,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 32, no. 1, pp. 1–4, 1990.
- [73] S. Yuferev and N. Ida, “Efficient implementation of the time domain surface impedance boundary conditions for the boundary element method,” *IEEE transactions on magnetics*, vol. 34, no. 5, pp. 2763–2766, 1998.
- [74] Q. Chen, M. Lu, and E. Michielssen, “Integral-equation-based analysis of transient scattering from surfaces with an impedance boundary condition,” *Microwave and Optical Technology Letters*, vol. 42, no. 3, pp. 213–220, 2004.
- [75] S. A. Sauter and M. Schanz, “Convolution quadrature for the wave equation with impedance boundary conditions,” *Journal of Computational Physics*, vol. 334, pp. 442–459, 2017.

- [76] L. Banjai and A. Rieder, “Convolution quadrature for the wave equation with a nonlinear impedance boundary condition,” *Mathematics of Computation*, vol. 87, no. 312, pp. 1783–1819, 2018.
- [77] A. E. Yilmaz, D. S. Weile, B. Shanker, J.-M. Jin, and E. Michielssen, “Fast analysis of transient scattering in lossy media,” *IEEE Antennas and Wireless Propagation Letters*, vol. 1, no. 1, pp. 14–17, 2002.
- [78] D. Dunavant, “High degree efficient symmetrical gaussian quadrature rules for the triangle,” *International Journal for Numerical Methods in Engineering*, vol. 21, no. 6, pp. 1129–1148, 1985.
- [79] M. G. Duffy, “Quadrature over a pyramid or cube of integrands with a singularity at a vertex,” *SIAM journal on Numerical Analysis*, vol. 19, no. 6, pp. 1260–1262, 1982.
- [80] D. Wilton, S. Rao, A. Glisson, D. Schaubert, O. Al-Bundak, and C. Butler, “Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains,” *IEEE Transactions on Antennas and Propagation*, vol. 32, no. 3, pp. 276–281, 1984.
- [81] A. Tzoulis and T. Eibert, “Review of singular potential integrals for method of moments solutions of surface integral equations,” *Advances in Radio Science*, vol. 2, pp. 93–99, 2004.
- [82] S. A. Sauter and C. Schwab, *Boundary Element Methods*. Springer, 2010.
- [83] R. D. Graglia, “On the numerical integration of the linear shape functions times the 3-D green’s function or its gradient on a plane triangle,” *IEEE Transactions on Antennas and Propagation*, vol. 41, no. 10, pp. 1448–1455, 1993.
- [84] W. C. Gibson, *The method of moments in electromagnetics*. Chapman and Hall/CRC, 2007.

Titre : Stratégies Computationnelles pour des Equations Intégrales avec Conditions d'Impédance aux Frontières en Domaines Fréquentiel et Temporel.

Mots clés : Equation Intégrale du Champ Electrique (EFIE), Conditions d'Impédance aux Frontières (IBC), Préconditionnement, Domaine Fréquentiel, Domaine Temporel

Résumé : L'équation intégrale du champ électrique (EFIE) est très utilisée pour résoudre des problèmes de diffusion d'ondes électromagnétiques grâce à la méthode aux éléments de frontière (BEM). En domaine fréquentiel, les systèmes matriciels émergeant de la BEM souffrent, entre autres, de deux problèmes de mauvais conditionnement : l'augmentation du nombre d'inconnues et la diminution de la fréquence entraînent l'accroissement du nombre de conditionnement. En conséquence, les solveurs itératifs requièrent plus d'itérations pour converger vers la solution, voire ne convergent pas du tout. En domaine temporel, ces problèmes sont également présents, en plus de l'instabilité DC qui entraîne une solution erronée en fin de simulation. La discrétisation en temps est obtenue grâce à une quadrature de convolution basée sur les méthodes de Runge-Kutta implicites.

Dans cette thèse, diverses formulations d'équations intégrales utilisant notamment des conditions d'impédance aux frontières (IBC) sont étudiées et préconditionnées. Dans une première partie en domaine fréquentiel, l'IBC-EFIE est stabilisée pour les basses fréquences et les maillages denses grâce aux projecteurs quasi-Helmholtz et à un préconditionnement de type Calderón. Puis une nouvelle forme d'IBC est introduite, ce qui permet la construction d'un préconditionneur multiplicatif. Dans la seconde partie en domaine temporel, l'EFIE est d'abord régularisée pour le cas d'un conducteur électrique parfait (PEC), la rendant stable pour les pas de temps larges et immunisée à l'instabilité DC. Enfin, une résolution efficace de l'IBC-EFIE est recherchée, avant de stabiliser l'équation pour les pas de temps larges et les maillages denses.

Title: Computational Strategies for Impedance Boundary Condition Integral Equations in Frequency and Time Domains

Keywords: Electric Field Integral Equation (EFIE), Impedance Boundary Condition (IBC), Preconditioning, Frequency Domain, Time Domain

Abstract: The Electric Field Integral Equation (EFIE) is widely used to solve wave scattering problems in electromagnetics using the Boundary Element Method (BEM). In frequency domain, the linear systems stemming from the BEM suffer, amongst others, from two ill-conditioning problems: the low frequency breakdown and the dense mesh breakdown. Consequently, the iterative solvers require more iterations to converge to the solution, or they do not converge at all in the worst cases. These breakdowns are also present in time domain, in addition to the DC instability which causes the solution to be completely wrong in the late time steps of the simulations. The time discretization is achieved using a convolution quadrature based on Implicit Runge-Kutta (IRK) methods, which yields a system that is solved by Marching-On-in-Time (MOT).

In this thesis, several integral equations formulations, involving Impedance Boundary Conditions (IBC) for most of them, are derived and subsequently preconditioned. In a first part dedicated to the frequency domain, the IBC-EFIE is stabilized for the low frequency and dense meshes by leveraging the quasi-Helmholtz projectors and a Calderón-like preconditioning. Then, a new IBC is introduced to enable the development of a multiplicative preconditioner for the new IBC-EFIE. In the second part on time domain, the EFIE is regularized for the Perfect Electric Conductor (PEC) case, to make it stable in the large time step regime and immune to the DC instability. Finally, the solution of the time domain IBC-EFIE is investigated by developing an efficient solution scheme and by stabilizing the equation for large time steps and dense meshes.