Doctoral Thesis

3-Dimensional Eigenmodal Analysis of Electromagnetic Structures

Author(s):
Guo, Hua

Publication Date:
2012

Permanent Link:
https://doi.org/10.3929/ethz-a-007599690

Rights / License:
In Copyright - Non-Commercial Use Permitted
3-dimensional eigenmodal analysis of electromagnetic structures

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
Hua Guo

MSc in Engineering, Huazhong University of Science and Technology
born May 1st, 1984
citizen of People’s Republic of China

accepted on the recommendation of
Prof. Dr. Peter Arbenz (ETH Zurich), examiner
Dr. Benedikt Oswald (Paul Scherrer Institute), co-examiner
Prof. Dr. Olivier Martin (EPF Lausanne), co-examiner
Prof. Dr. Lukas Novotny (ETH Zurich), co-examiner

2012
Acknowledgements

This work could not have been completed without the help of a number of people. First, I would like to express my sincere gratitude to Prof. Dr. Peter Arbenz, main supervisor and Dr. Benedikt Oswald, co-supervisor. Peter Arbenz generously gave me the opportunity to undertake my doctoral research project in his group. He has been a fantastic teacher and colleague. He lent me tremendous support for solving numerable delicate problems. Also, he granted considerable freedom to study my own ideas. I have profited enormously from his profound knowledge and patient guidance over all these years.

Another special thanks goes to Benedikt Oswald. He has always encouraged me to investigate challenging problems, in order to broaden my scientific horizon. He has always been here to answer questions, and share ambitious scientific concepts with me. This work would never have come into existence without his contribution.

I am very much obliged to Prof. Olivier Martin, co-referee of this thesis, who has provided many thoughtful and supportive reviews and suggestions with respect to innovative plasmonic nanostructures. Also, I would like to thank Prof. Lukas Novotny, who kindly agreed to co-examine this thesis.

Furthermore, I want to thank the following people: Dr. Andreas Kern for sharing his experience on nano-antennas; Dr. Arya Fallahi for teaching me theoretical concepts for modeling nano-spheres; Dr. Andreas Adelmann, Dr. Antonio Falone and Dr. Gian-Luca Orlandi for providing me the geometry of the transverse deflecting cavity and giving me insight on its design principles.

I would also like to thank all current and former colleagues in the group: Yves Ineichen, Christof Kraus, Dr. Cyril Flaig, Dr. Erhan Turan, Marcus Wittberger, Stefan Pauli and Dr. Roman Geus.

This work was sponsored in part by grant no. 200021-117978 of the Swiss National Science Foundation (SNF).

All simulations were carried out on the infrastructure of the Swiss National Supercomputing Centre (CSCS), which is herewith gratefully acknowledged.

Last but not least, I wish to thank all those who helped me to overcome many difficult situations. I am especially grateful to my parents, my wife Yuanrui Zhang, and of course all my friends for their constant solid support during this exciting period.
Summary

In this project the eigensolver package Femaxx was substantially extended so that it is now capable of calculating electromagnetic eigenmodes including almost all physically reasonable loss mechanisms. Before, Femaxx had been developed to compute electromagnetic eigenmodes in closed cavities, bounded by perfect electric conductor walls. Femaxx generally discretizes the time-harmonic Maxwell Equations with the finite element method (FEM) in 3-dimensional space on unstructured tetrahedral grids, in order to model complicated curved geometry. The code has been parallelized and optimized for distributed memory parallel computing architectures.

Femaxx implements the Jacobi-Davidson QZ (JDQZ) method, for solving linear or quadratic eigenvalue problems. In this thesis, we have used the JDQZ eigensolver for the analysis of a wide span of resonant electromagnetic structures. First, we analyze several bounded, resonant cavities, explicitly including electromagnetic loss mechanisms in the model. We compare our numerical results with published data, and benchmark our new method. We also demonstrate the performance and parallel efficiency of the JDQZ eigensolver. Second, we calculate the modes of several dielectric resonator antennas (DRA), and compare them to theoretically derived results. Femaxx also implements transparent boundary conditions that enable the analysis of DRA in unbounded space. Third, we analyze several plasmonic nanostructures in unbounded space through solving a sequence of quadratic eigenvalue problems. Both dissipation and radiation loss mechanisms are considered in the model.

In addition, Femaxx implements the nonlinear Jacobi-Davidson (NLJD) method. Using a fully iterative scheme of NLJD eigensolvers, we analyze plasmonic nanostructures, fully considering the dispersive dielectric properties of metals in the optical region of the electromagnetic spectrum. We investigate the properties of the algorithm in detail and demonstrate its performance by analyzing a specific type of an optical antenna.

The nonlinear Jacobi-Davidson eigensolver also solves the nonlinear eigenvalue problem that arises from the integration of the skin effect into a model for enclosing cavity walls with finite ohmic conductivity. This assumption then leads to the concept of surface impedance boundary conditions (SIBC) which collapse a electromagnetically lossy volume region into a 2-dimensional sheet.

Femaxx package also contains a post-processor, which calculates the resonance, loss, quality factor and visualizes the electromagnetic field distribution, associated with a specific eigenmode. We are confident that Femaxx has become a useful tool in micro- and nano-fabrication practice.
Zusammenfassung

# Contents

1 Introduction .............................................. 1  
1.1 Overview .............................................. 1  
1.2 Review of the original lossless Femaxx version .............. 2  
  1.2.1 Formulation of the problem ......................... 2  
  1.2.2 Numerical experiments ............................... 5  
1.3 Concluding remark .................................... 9  

2 Electromagnetic eigenvalue problem with loss .................. 11  
2.1 Introduction .......................................... 11  
2.2 Loss mechanisms .................................... 11  
2.3 The JDQZ eigensolver .................................. 12  
2.4 Complex-valued dielectric material properties ............... 17  
  2.4.1 Formulation of the problem ....................... 17  
  2.4.2 Numerical Experiments .......................... 19  
2.5 Ohmic conductivity loss in volume regions .................. 21  
  2.5.1 Formulation of the quadratic eigenvalue problem ...... 21  
  2.5.2 Linearization of the quadratic eigenvalue problem (QEP) 24  
  2.5.3 Numerical experiments ....................... 25  
2.6 Summary .............................................. 28  

3 Plasmonic nanostructures .................................. 29  
3.1 Introduction ........................................... 29  
3.2 Formulation of the problem ............................ 31  
3.3 Numerical method .................................... 34  
  3.3.1 The finite element method .................. 34  
  3.3.2 The eigensolver ................................ 36  
3.4 Validation and application of the algorithm ................ 37  
  3.4.1 Dielectric resonator antenna ................... 38  
  3.4.2 Cuboid ........................................... 41  
  3.4.3 Sphere ......................................... 42  
  3.4.4 An optical dipole antenna ....................... 46  
3.5 Discussion and conclusions ................................ 51  

4 A nonlinear eigensolver for the analysis of dispersive nanostructures 55  
4.1 Introduction ........................................... 55  
4.2 Formulation of the problem ............................ 57  
4.3 Nonlinear eigensolver .................................... 60
Contents

4.4 Numerical experiments ................................................. 65
  4.4.1 A ‘small’ nanoantenna ............................................ 67
  4.4.2 A ‘large’ nanoantenna ............................................ 72
4.5 Discussions and Conclusions ........................................... 74

5 Electromagnetic eigenmodal analysis of a molecular nanosensor 77
  5.1 Introduction .......................................................... 77
  5.2 Numerical experiment ............................................... 78
  5.3 Summary ............................................................. 81

6 Cavity eigenmodal analysis with surface impedance boundary condi-
tions ................................................................. 83
  6.1 Formulation of the problem .......................................... 83
  6.2 Nonlinear eigensolver .............................................. 87
  6.3 Numerical experiments ............................................ 88
    6.3.1 Pillbox cavity .................................................. 88
    6.3.2 Transverse deflecting cavity .................................. 89

7 Implementation ............................................................ 91
  7.1 Mesh ................................................................. 91
  7.2 Trilinos ............................................................ 92
  7.3 Postprocessing ..................................................... 93
  7.4 Object-oriented design ............................................ 93

8 Future work .................................................................. 95

A Femaxx User Guide ....................................................... 97
  A.1 femaxx_lep_cdriver ................................................ 97
  A.2 femaxx_qep_cdriver .............................................. 100
  A.3 femaxx_nlep_cdriver ............................................. 104
  A.4 femaxx_cpost ....................................................... 109

B Finite Element Method .................................................. 111

Bibliography ................................................................. 115

Curriculum Vitae ........................................................ 123
List of Figures

1.1 We show the geometry and the tetrahedral mesh of the 5-cell transverse deflecting cavity in (a). We plot the electric field distribution $E$, in a cut plane along the axis of the cavity, for the TM110 mode in (b). 6

1.2 We plot $E$ and $H$ along the z-axis for TM110 mode: (a) $E_x$ and $E_y$; (b) $H_x$ and $H_y$. Here, $E_z$ and $H_z$ ($\approx 0$) are not shown. 7

1.3 We plot the electric field distribution of for four different modes of the completed PSI cyclotron system consisting of a combination of four simplified cavities coupled to the beam tube. 8

2.1 We show the geometry of the lossy dielectric block cavity in (a). We give a cross-sectional view in (b) where the light gray area represents vacuum material properties while the gray area represents the lossy dielectric block. 19

2.2 We show the electromagnetic field distribution of the cavity which contains a lossy dielectric block; in particular, we show the two lowest-order modes, visualized in the plane where the lossy dielectric block is attached to the metallic wall, cf. Fig. 2.1. 22

2.3 We show the electromagnetic field distribution of the dominant mode in a rectangular resonator whose right half is filled with ohmically lossy material: $\varepsilon_r = 2.0$ and $\sigma = 1.3$ (S/m). 26

2.4 We show the geometry of the cylindrical cavity with circular cross-section containing in turn an ohmically conducting rod, marked by the red area. 27

2.5 We show the electric field distributions for the cylindrical cavity containing a conductive rod with circular cross-section: (a) mode 1 and (b) mode 4. 28

3.1 A 3-D view of a typical computational domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$. $\Omega_1$ is the optical device, $\Omega_2$ is the substrate, $\Omega_3$ is the environment. $\Gamma_{13}$ is the interface between $\Omega_1$ and $\Omega_2 \cup \Omega_3$, and $\Gamma_{23}$ is the interface between $\Omega_2$ and $\Omega_1 \cup \Omega_3$. $\Gamma$ is the boundary of $\Omega$. 32
3.2 (a) a 3-D sketch of the DRA (or nano-cuboid). (b): a DRA with $a = b = 10\text{ mm}$, $d = 4\text{ mm}$ and $\varepsilon_r = 20.0$, resonance at 6.545 GHz. The electric field distribution $|\mathbf{E}|$ of the $TE^{x}_{111}$ mode in the antenna and its surrounding region is shown. (c): a gold cuboid with $a = b = 100\text{ nm}$, $d = 40\text{ nm}$, resonance at 935.1 THz (where $\varepsilon_r = -1.2308 + i5.8458$). It shows the electric field distribution $|\mathbf{E}|$ of the $TE^{x}_{111}$ mode in the cuboid and the surrounding region. (b) and (c) are visualized on the $xy$-plane through the center of the antenna. (d) 3 components of the electric field of the DRA (in (b)) plotted along the $y$-axis. (e) 3 components of the electric field of the gold cuboid (in (c)) plotted along the $y$-axis. Note that, $E_x$, $E_y$, and $E_z$ are complex. For $E_y$ and $E_z$ (approximately zero), only real parts are shown. $E_x$ is scaled such that $\|\text{Re}(E_x)\|_{\infty} = 1$. 40

3.3 Mie solutions of silver spheres with radii of (a) 30 nm and (b) 60 nm. 43

3.4 Numerical and Mie solutions of Q factors and radiative quantum yield (subfigure) of silver spheres with varying radius $R$. 45

3.5 (a) and (b) show the electric field distributions ($|\mathbf{E}|$) in the vicinity of a silver sphere on the $xy$-plane. The radius of the sphere is 60 nm. (a): the mode when $Q_{\text{sca}}$ reaches maximum, $E_{\text{res}} = 3.03\text{ eV}$; (b): the mode when $Q_{\text{abs}}$ reaches maximum, $E_{\text{res}} = 3.48\text{ eV}$. 45

3.6 The nano-optical dipole antenna. (a) geometry; (b) 3-D sketch. 46

3.7 (Numerical analysis of a gold, nano-optical dipole antenna (model (1)). The gap width is 20 nm and the resonance is at 665.7 nm; (a) magnetic and (b) electric field distribution ($|\mathbf{H}|$ and $|\mathbf{E}|$) in the vicinity of the antenna surface; (c) electric field distribution $|\mathbf{E}|$ around the gap; (a), (b), and (c) are evaluated on the $xy$-plane through the center of the antenna arms. 47

3.8 plots the electric field amplitudes $|\mathbf{E}|$ along the $x$-axis, associated with 3 material arrangements. 48

3.9 Numerical analysis of a gold nano-optical antenna (model (2)): (a) gap width $= 5\text{ nm}$, resonance at 796.3 nm. It shows the electric field distribution $|\mathbf{E}|$ (visualized on the $xy$-plane) in the vicinity of the antenna surface; (b) plots of the electric field amplitude $|\mathbf{E}|$ along the $x$-axis with varying gap width $g$. 49
3.10 (a) Charge profiles; (b) the dark mode of the gold nano-optical antenna (gap width $g = 10$ nm) in vacuum: resonance at 557.4 nm. The electric field distribution $|E|$ (visualized on the $xy$-plane) in the vicinity of the antenna surface is shown; (c) the gold nano-optical antenna (gap width $g = 5$ nm) in vacuum. The electric field amplitude $|E|$ along the $x$-axis is plotted, associated with the bright mode (resonance at 713.9 nm) and the dark mode (resonance at 532.1 nm), respectively.

4.1 A 3-D view of a typical computational domain $\Omega = \Omega_1 \cup \Omega_2$. $\Omega_1$ is the optical device, $\Omega_2$ is the surrounding medium, $\Gamma_{12}$ is the interface between $\Omega_1$ and $\Omega_2$, $\Gamma$ is the boundary of $\Omega$. In Section 4.4, we let $\Omega$ be a sphere in order to reduce the error caused by the 1st order absorbing boundary condition [29].

4.2 The relative permittivity $\varepsilon_r$ for gold as a function of the photon energy [40].

4.3 Plot of the parameter $D$.

4.4 (a) Cross sectional view and (b) a 3-D sketch of the nanoantenna with cylindrical disks.

4.5 The convergence history for the fine mesh. The $y$ axis denotes the normalized residual $\|r\|_2$, and the $x$ axis denotes the NLJD iteration step (IT). Linear (a) and quadratic (b) elements are used.

4.6 Distribution of the eigenmodes for the gold nanoantenna within the infrared range $S_{\text{IR}}$; (+) radius $r = 60$ nm; (●) radius $r = 120$ nm. The results are obtained by quadratic elements.

4.7 The electric field distribution $|E|$ in the gap of gold nanoantennas obtained with quadratic elements. (a)–(d) radius $r = 60$ nm; (e)–(k) radius $r = 120$ nm. (a) is visualized on the $yz$-plane, and all the other sub-figures are visualized on the $xy$-plane through the center of the gap.

5.1 (a) Geometry of the dipole antenna. (b) The sketch for molecular nonsensing: both the top view (left) and the cross section (right) are shown. The yellow area is the optical dipole antenna, and it resides on a silica substrate (gray area) and it is surrounded by a molecule layer (light gray area).

5.2 The resonance $\lambda_{\text{res}}$ of the bright mode vs. the refractive index $n$ of the molecule layer. The thickness of the layer is $t = 5$ nm.
List of Figures

5.3  (a) The sensitivity (nm/RIU) of the nanosensor vs. the thickness of the molecule layer $t$ (nm). (b) Electric field distribution $|E|$ in the vicinity of the antenna surface. The layer has $n = 1.5$ and $t = 40$ nm. The resonant wavelength is $\lambda_{res} = 924$ nm. (b) is evaluated on the $xy$-plane through the center of the antenna arms. .................................................. 80

6.1  We show the electric field distribution $|E|$ for TM010 mode of a pillbox cavity with surface conductivity $\sigma_s = 5.8 \cdot 10^7$ S/m, visualized in the half domain. ................................. 89
## List of Tables

2.1 Numerical results of the lossy dielectric block cavity: linear elements are used. \( f \): the resonant frequency \( = \omega/(2\pi) \); \( Q_1 \): the quality factor defined by Eq. (2.12); \( Q_2 \): the quality factor defined by Eq. (2.17). ................................................................. 20

2.2 Numerical results of the lossy dielectric block cavity: quadratic elements are used. \( f \): the resonant frequency \( = \omega/(2\pi) \); \( Q_1 \): the quality factor defined by Eq. (2.12); \( Q_2 \): the quality factor defined by Eq. (2.17). ................................................................. 20

2.3 We show the parallel efficiency \( E(p) \) of the JDQZ eigensolver. Experiments are carried out for quadratic elements on Mesh3 (#dof = 855'912) ................................................................. 21

2.4 The numerical results the dominant mode of the half-filled rectangular resonator. \( \tilde{\omega} \): the complex angular frequency (Unit: GH兹), \( \tilde{\omega} = \omega + i\alpha \); \( Q_1 \): the quality factor defined by Eq. (2.12); \( Q_2 \): the quality factor defined by Eq. (2.17). ................................................................. 25

2.5 We show the numerical results of the five lowest-order eigenmodes for the cylindrical cavity containing a conductive rod with circular cross-section. \( \tilde{\omega} \): the complex angular frequency (Unit: MHz), \( \tilde{\omega} = \omega + i\alpha \); \( Q_1 \): the quality factor defined by Eq. (2.12); \( Q_2 \): the quality factor defined by Eq. (2.17). ................................................................. 27

3.1 \( TE_{111}^z \) mode of DRAs \( (f \): resonant frequency computed by \texttt{Femmaxx}; \( f_{\text{MI}} \): theoretically determined resonant frequency [50]; Diff.: relative difference of \( f \) and \( f_{\text{MI}} \); \( Q_{\text{MI}} \): theoretically determined the Q factor [50].) ................................................................. 39

3.2 Numerical figures of merit for the \( TE_{111}^z \) mode of the nanocuboid where \( \lambda \): resonant wavelength computed by \texttt{Femmaxx}; \( f \): resonant frequency computed by \texttt{Femmaxx}; \( f_{\text{MI}} \): theoretically determined resonant frequency [50]; Diff.: relative difference of \( f \) and \( f_{\text{MI}} \); \( Q_{\text{MI}} \): theoretically determined Q factor [50]. ................................................................. 42

3.3 Results of the numerical analysis and the semi-analytical Mie solutions of the silver sphere hovering in vacuum, where: \( E_{\text{Res1}} \): resonant energy computed by \texttt{Femmaxx}; \( E_{\text{Diff}} \): the relative difference between \( E_{\text{Res1}} \) and \( E_{\text{Mie}} \); \( Q_{\text{Diff}} \): the relative difference between \( Q_1 \) and \( Q_{\text{Mie}} \). ................................................................. 44
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Numerical analysis of a gold nano-optical antenna (#p: number of cores; N: number of tetrahedra; dofs: degrees of freedom; λ: resonant wavelength computed by femaxx; Q₁ and Q₂: quality factors by Eq. (3.5) and Eq. (3.10), respectively)</td>
<td>48</td>
</tr>
<tr>
<td>3.5</td>
<td>Dark modes vs. bright modes</td>
<td>50</td>
</tr>
<tr>
<td>4.1</td>
<td>Numerical results for the small nanoantenna (r = 60 nm) and a fine mesh with 436’691 tetrahedra.</td>
<td>68</td>
</tr>
<tr>
<td>4.2</td>
<td>Numeric results for the small nanoantenna (r = 60 nm) and a coarse mesh with 71’003 tetrahedra.</td>
<td>69</td>
</tr>
<tr>
<td>4.3</td>
<td>Numerical analysis of a gold nanoantenna with r = 120 nm. Here the quadratic elements are employed.</td>
<td>72</td>
</tr>
<tr>
<td>6.1</td>
<td>Numerical analysis for TM010 mode of the pillbox cavity. f = ( \frac{\omega}{2\pi} ) is the resonance frequency; ( \delta ) is the skin depth (6.5); Q is the quality factor (6.11).</td>
<td>88</td>
</tr>
</tbody>
</table>
Chapter

Introduction

1.1 Overview

The numerical computation of eigenfrequencies and corresponding eigenmodal fields of various electromagnetic structures, based on full-wave, 3-dimensional models, has attracted increasing interest in recent years [29, 85, 71, 69]. Since the manufacturing process is expensive and time-consuming, a device that does not comply with its original performance specification is not acceptable. Therefore, reliable and accurate pre-fabrication computer simulation is indispensable. Since 2002, in a collaboration between ETH Zurich and Paul Scherrer Institut, the software package Femaxx has been continuously developed.

The first Femaxx version was released in 2006 [23, 3]. This version solves the 3-dimensional, time-harmonic electric field vector wave, aka. curl-curl, equation. In particular, this version does not consider any loss mechanisms, which is perfectly acceptable when analyzing resonant cavities bounded by metallic walls with negligible ohmic loss. It generally solves large scale linear eigenvalue problems on distributed memory computing architectures.

Femaxx discretizes the electric curl-curl equation with the finite element method (FEM) on unstructured, tetrahedral grids in order to model geometrically complicated geometry with curved boundaries. In addition using the finite element method, it becomes possible to efficiently discretize geometry with a wide span of characteristic length scales, i.e., modeling with level of detail, using small elements for delicate geometrical features and larger elements elsewhere.

Femaxx employs linear or quadratic Nédélec elements [39], which are mandatory in order to represent the electro-magnetic field vector. The Nédélec elements generally eliminate the shortcomings of node-based finite elements in electrodynamics calculations. In particular, using Nédélec elements avoids spurious solutions. Furthermore, since the Nédélec elements are designed in such a way that the tangential components of the electro-magnetic field are continuous across elemental boundaries, imposing boundary conditions is straight-
1 Introduction

forward. In the first version, we use the perfect electric conductor (PEC) boundary condition, which implies that the tangential electric field identically vanishes on the cavity wall.

Within the framework of this thesis, starting in 2008, Femaxx was extended, going considerably beyond the original version. Femaxx can now analyze resonant electromagnetic structures including all relevant electromagnetic loss mechanisms. Additionally, it is now possible to model resonating structures in free space. Here, we list the physical models which were implemented into the new Femaxx version.

1. imperfect dielectric materials, i.e., dielectric loss in a volume region, c.f. chapter 2.

2. ohmic conductors, i.e., finite ohmic conductivity loss in a volume region, c.f. chapter 2.

3. finite sized electromagnetic structures resonating in unbounded space; e.g. a dielectric resonator antenna (DRA) radiating into free space, c.f. chapter 3.

4. highly dispersive plasmonic nanostructures, e.g., optical dipole antenna, c.f. chapters 3, 4 and 5.

5. finite ohmic conductivity in a 2-dimensional sheet region, e.g., cavity wall conductivity modeled with the frequency-dependent skin effect, c.f. chapter 6.

The inclusion of the above physical mechanisms has required the implementation of several Jacobi-Davidson type eigensolvers, in order to solve complex linear, quadratic and even nonlinear eigenvalue problems. All solvers have been parallelized on distributed memory computing architectures. In this thesis, we also study the algorithms and their respective performance.

1.2 Review of the original lossless Femaxx version

We employ the original Femaxx for studying resonant accelerator cavities bounded by perfect electric conductor (PEC) walls [60].

1.2.1 Formulation of the problem

Let \( \Omega \) be a closed domain describing an resonant cavity structure. Femaxx-1.0 solves the electric field curl-curl equation in \( \Omega \) in the time-harmonic regime. We assume that no external fields, sources or charges are present and we
1.2 Review of the original lossless Femaxx version

also exclude loss mechanisms. Eventually we obtain the following form of the Maxwell equation [23]

\[ \nabla \times (\mu_r^{-1} \nabla \times E(x)) - k_0^2 \varepsilon_r E(x) = 0, \quad x \in \Omega, \quad (1.1a) \]

with the divergence-free condition

\[ \nabla \cdot (\varepsilon_r E(x)) = 0, \quad x \in \Omega, \quad (1.1b) \]

where \( \varepsilon_r \) and \( \mu_r \) are relative magnetic permittivity and relative electric permeability, respectively. In the absence of losses, \( \varepsilon_r \) and \( \mu_r \) are both real constants in \( \Omega \). On \( \Omega \)'s boundary, \( \Gamma = \partial \Omega \), the perfect electric conductor (PEC) boundary condition applies

\[ n \times E(x) = 0, \quad x \in \Gamma. \quad (1.1c) \]

A natural weak form of (1.1) [79] is then

Find \( k_0 \in \mathbb{C} \) and \( E \in V \), \( E \neq 0 \), such that for all \( f \in V \) and all \( q \in W \)

\[ \int_{\Omega} \mu_r^{-1} \nabla \times f \cdot \nabla \times E \, dx - k_0^2 \int_{\Omega} \varepsilon_r f \cdot E \, dx = 0, \]

\[ \int_{\Omega} \varepsilon_r E \cdot \nabla q \, dx = 0. \quad (1.2) \]

Here, \( V \) denotes the functions in \( H(\text{curl}; \Omega) \) that satisfy the PEC (1.1c) and \( W = H_0^1(\Omega) \). Details on these function spaces are given in [39].

To numerically evaluate the natural weak form Eq. (1.2), we decompose the computational domain \( \Omega \) into tetrahedra and we discretize (1.2) with the Ritz-Galerkin method [39] using finite element subspaces of \( V \) and \( W \). The electric vector functions in \( V \) are approximated by Nédélec elements, while the scalar functions in \( W \) are approximated by Lagrange, or nodal, elements [39]. This approach avoids spurious solutions. Imposing the PEC boundary condition is also straightforward [39].

We let the vector functions \( N_i \), \( 1 \leq i \leq n \), be the Nédélec basis functions, while the scalar functions \( N_\ell \), \( 1 \leq \ell \leq m \), denote Lagrange basis functions. Then, the weak form (1.2) leads to a real-valued, linear, generalized matrix eigenvalue problem [23]

\[ A x = \lambda M x, \quad (1.3) \]

with a constraint arising from the divergence-free condition

\[ C^T x = 0, \quad (1.4) \]

where \( \lambda(= k_0^2) \) is the eigenvalue and \( x \) is the eigenvector. The matrices \( A, M \in \mathbb{R}^{n \times n} \).
1 Introduction

$\mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{n \times m}$ in Eqs. (1.3) and (1.4) have entries

\[
\begin{align*}
a_{ij} &= \int_{\Omega} \mu_r^{-1} (\nabla \times N_i) \cdot (\nabla \times N_j) \, dx, \quad 1 \leq i, j \leq n, \\
m_{ij} &= \int_{\Omega} \varepsilon r N_i \cdot N_j \, dx, \quad 1 \leq i, j \leq n, \\
c_{i\ell} &= \int_{\Omega} \varepsilon r N_i \cdot \nabla N_\ell \, dx, \quad 1 \leq i \leq n, \quad 1 \leq \ell \leq m.
\end{align*}
\]

Both $A$ and $M$ are real symmetric; $M$ is positive definite, and $A$ is positive semi-definite. In the original Femaxx version, the Jacobi-Davidson symmetric type (JDSYM) eigensolver was implemented to compute eigenpairs of Eqs. (1.3) and (1.4). For details of algorithms and performance of JDSYM we refer readers to [23, 3].

The dimension of $A$’s null space is significant and the vectors in the null space of $A$ correspond to $\mathbf{E}(x)$ having non-zero divergence [23], thus violating the divergence free condition, Eq. (1.4). By explicitly imposing (1.4), Femaxx avoids computing zero eigenvalues.

In order to compute the entries $c_{i\ell}$ of $C$, we need to compute the gradient $\nabla N_\ell$ of the Lagrange basis function $N_\ell$. We note that $\nabla N_\ell$ associated with the mesh node $\ell$ is a linear combination of the Nédélec basis functions associated with the edges connected to node $\ell$. The linear factors are either $-1$ or $1$, depending on the direction of the edge. (The factor is $1$ if the edge points to the node $\ell$, and $-1$ if it points away from it.) Similarly, the gradient of the (quadratic) Lagrange basis function associated with a particular edge midpoint is equal to the quadratic Nédélec basis function associated with that edge. Therefore [23, 25],

\[
\nabla N_\ell = \sum_{j=1}^{m} y_{j\ell} N_j, \quad 1 \leq \ell \leq m,
\]

where the coefficients $y_{j\ell}$ are either 1, 0, or $-1$ with at most two being nonzero. By inserting (1.6) into the definition of $C$ and considering the definition of $M$ in (1.5), it is obvious that the columns of $C$ are linear combinations of the columns of $M$, i.e.,

\[
C = M[y_1, \ldots, y_m] = MY, \quad y_{\ell} = (y_{1\ell}, \ldots, y_{n\ell})^T,
\]

where the components $y_{j\ell}$ form a sparse rectangular matrix $Y \in \mathbb{R}^{n \times m}$. $Y$ can be easily built from the mesh data.

The desired eigensolutions $(\lambda, \mathbf{x})$ satisfy $Y^T M \mathbf{x} = C^T \mathbf{x} = 0$. We use a
1.2 Review of the original lossless Femaxx version

M-orthogonal projector $P$ [23]

$$ P = I - Y H^{-1} C^T, \quad H = Y^T M Y = Y^T C, \quad (1.8) $$

to keep the search space in the null space of $C^T$. This implies that all search vectors satisfy the divergence-free condition. The projector $P$ is applied to a vector as

$$ P x = x - Y (H^{-1} (C^T x)). \quad (1.9) $$

1.2.2 Numerical experiments

To demonstrate Femaxx’s capabilities, we study two different resonant accelerator cavities, fabricated at PSI. We use quadratic Nédélec elements, where each element has 20 local degrees of freedom (dof). In order to accurately model the complicated geometry, c.f. Fig 1.1(a), we use highly resolved tetrahedral meshes. Since the size of the eigenvalue problem is usually large, meaning millions of degrees of freedom, all simulations were run on the Cray of the Swiss National Supercomputing Centre (CSCS) [14].

Transverse deflecting cavity

The next generation X-ray Free Electron Laser (SwissFEL) under development at PSI will employ a 5-cell transverse deflecting cavity (TDC) [60, 26, 17] for beam diagnostics. Each cell is a pillbox, and the middle cell is coupled to a rectangular feeding waveguide, Fig. 1.1(a). We assume vacuum material properties inside the cavity, i.e., $\varepsilon_r = \mu_r = 1$. While the cavity boundary is modeled with copper having conductivity $\sigma_s = 5.8 \times 10^7$ S/m. Since $\sigma_s \gg 1$, using the PEC boundary condition, Eq. (1.1c), is justified [63]. The cavity’s details are given in [60].

We use the Gmsh tetrahedral mesh generation program [24] to generate a mesh containing approximately 870’000 tetrahedra, Fig. 1.1(a). This results more than 5 million degrees of freedom. The calculated resonance frequency of the operating TM110 mode is 2.995930 GHz. This matches with the design parameters [60], given as 2.997912 GHz. The electric field distribution is shown in Fig 1.1(b).

We denote the 3 components of the electric field $E$ and magnetic field $H$ along the z-axis, through the center of the cells, with $E_x$, $E_y$, $E_z$ and $H_x$, $H_y$, $H_z$. For the TM110 mode, we obtain $E_y \approx 0$, $E_z \approx 0$, and $E_x \neq 0$, Fig. 1.2(a), similarly, $H_x \approx 0$, $H_z \approx 0$, and $H_y \neq 0$, Fig. 1.2(b).
1 Introduction

Figure 1.1: We show the geometry and the tetrahedral mesh of the 5-cell transverse deflecting cavity in (a). We plot the electric field distribution $\mathbf{E}$, in a cut plane along the axis of the cavity, for the TM110 mode in (b).
1.2 Review of the original lossless *Femaxx* version

Figure 1.2: We plot $E$ and $H$ along the $z$-axis for TM110 mode: (a) $E_x$ and $E_y$; (b) $H_x$ and $H_y$. Here, $E_z$ and $H_z$ ($\approx 0$) are not shown.
1 Introduction

Cyclotron

We calculate the electromagnetic eigenmodes that exist inside the geometry of the complete PSI cyclotron. The PSI cyclotron is modeled through the combination of four simplified cavities coupled to the beam tube [73], Fig 1.3. This system is characterized through enormous geometrical complexity. Therefore it

\begin{align*}
\text{(a) 1st mode: } & f_{res.} = 51.48 \text{ MHz} \\
\text{(b) 31st mode: } & f_{res.} = 101.04 \text{ MHz} \\
\text{(c) 86th mode: } & f_{res.} = 149.58 \text{ MHz} \\
\text{(d) 87th mode: } & f_{res.} = 151.51 \text{ MHz}
\end{align*}

Figure 1.3: We plot the electric field distribution of for four different modes of the completed PSI cyclotron system consisting of a combination of four simplified cavities coupled to the beam tube.

is an attractive test case for Femaxx. We again assume vacuum material properties, analogously to the case of the 5-cell TDC, c.f. section 1.2.2, and also model the cavity walls with a perfect electric conductor boundary. We generate a high-quality tetrahedral mesh containing approximately 650'000 elements with Gmsh [24]. We compute the first 100 electromagnetic modes successfully and show four of them in Figure 1.3. The computed resonant frequency of the fundamental (first) mode is 51.48 MHz while the associated measured value is 50.633 MHz [73]. This difference between computed and measured resonance
1.3 Concluding remark

frequencies originates from the difference between the simplified geometry of the simulated PSI cyclotron system and the real geometry [73].

1.3 Concluding remark

In the preceding two sections we have demonstrated the generation of efficient tetrahedral grids for realistic cavity geometries, calculated the respective eigenmodal electromagnetic fields and compared the numerical results to experimental data. Femaxx so far has been shown to be capable of solving large electromagnetic eigenvalue problems where no loss is involved and the boundary is implemented via perfect electric conductor walls.
2.1 Introduction

For the highly accurate design of modern accelerator cavities electromagnetic loss mechanisms must not be neglected anymore, since the eigenmodes may be affected considerably by loss. An electromagnetic model that neglects loss may serve well during the initial design stage, but is eventually insufficient in today’s manufacturing environment where fabrication cost must be minimized. Therefore loss must be generally included in the model. In this Chapter, we consider two different loss mechanisms, namely dielectric and ohmic loss, to study their effect onto the eigenmodal solutions.

We apply the PEC boundary condition with the finite element method, which then yields the generalized eigenvalue problem (1.3) with constraint (1.4). The matrix $M$, with entries depending on complex-valued $\varepsilon_r$ of the dielectric material, Eq (1.5), then becomes complex symmetric. Therefore Eq. (1.3) is a generalized, non-Hermitian eigenvalue problem (GNHEP). $C$ is also a complex rectangular matrix. In order to solve the generalized, non-Hermitian eigenvalue problem, we develop the Jacobi-Davidson QZ (JDQZ) method [68, 19], c.f. section 2.3.

When there is non-negligible ohmic loss, we derive a different version of the electromagnetic vector wave equation which includes ohmic current density, c.f. section 2.5. This leads to a quadratic eigenvalue problem which is also solved by the JDQZ method through linearization.

2.2 Loss mechanisms

While previous work [13, 85, 82] has considered loss in the solution of eigenmodal problems, it was generally restricted to small-sized eigenvalue problems, i.e., a few thousand degrees of freedom in the respective finite element
discretization. Here we restrict ourselves to isotropic, linear materials with either dielectric or ohmic loss. We neglect ferromagnetic materials. We assume $\mu_r = 1$ throughout this thesis. We then use the complex-valued relative permittivity given as

$$\varepsilon_r = \text{Re}(\varepsilon_r) - i \text{Im}(\varepsilon_r).$$

Here, the imaginary part $\text{Im}(\varepsilon_r)$ is not zero and directly relates to loss. In addition, when the material also contains free electrons or holes, we obtain the conduction current density

$$J = \sigma E,$$

here, $\sigma$ is the ohmic conductivity. The total current density [63] is then

$$J = i\omega\varepsilon_0(\text{Re}(\varepsilon_r) - i \text{Im}(\varepsilon_r) - i \frac{\sigma}{\varepsilon_0\omega})E.$$

In technically relevant dielectrics, only few free electrons or holes exist and the conduction current is included in $\text{Im}(\varepsilon_r)$, so $\sigma = 0$ [63]. On the other hand, in good conductors, conduction current dominates and bound electron effects are included in $\sigma$, so $\text{Im}(\varepsilon_r) = 0$ [63].

### 2.3 The JDQZ eigensolver

The Jacobi-Davidson QZ (JDQZ) algorithm 1 was taken from [68] and implemented with only a few modifications. Most of contents in this section are copied from [68].

With $\lambda = \zeta/\eta$, the generalized eigenproblem $Ax = \lambda Mx$ becomes

$$(\eta A - \zeta M)x = 0,$$  

and we denote the eigenvalue of the matrix pair $(A, M)$ as a pair of $(\zeta, \eta)$. The JDQZ algorithm finds the partial generalized Schur form of dimension $k$ for $A, M$ [68]

$$AQ_k = Z_k R_k^A, \quad MQ_k = Z_k R_k^M.$$  

Here, $Q_k$ and $Z_k$ are orthogonal $n \times k$ matrices; $R_k^A$ and $R_k^M$ are $k \times k$ upper triangular matrices. A column $q_i$ of $Q_k$ is the generalized Schur vector. The pair $((\zeta_i, \eta_i), q_i)$ with $(\zeta_i, \eta_i) = (R_k^A(i, i), R_k^M(i, i))$ is referred as the generalized Schur pair [68]. If $(\zeta, \eta)$ is the eigenvalue of $(R_k^A(i, i), R_k^M(i, i))$ and $y$ is the corresponding eigenvector, then $(\zeta/\eta, Qy)$ is the eigenpair of $(A, M)$, cf. steps 37–38 in Alg. 1.

Eq. (2.2) suggests

$$\eta_i Aq_i - \zeta_i Mq_i \perp z_i.$$
2.3 The JDQZ eigensolver

In each Jacobi-Davidson step, the approximate Schur vector \( \mathbf{u} \) is taken from a search subspace \( \text{span}(V) \). \( V \) is an orthonormal basis and the new search vector is orthogonalized w.r.t. the current basis by a modified Gram-Schmidt method (MGSM), cf. step 4. We will later discuss how to find the new search vector of \( V \) by solving the so-called correction equation. We employ the Petrov-Galerkin condition \([68]\), i.e. the corresponding residual is orthogonal to a test subspace \( \text{span}(W) \):

\[
\eta A \mathbf{u} - \zeta M \mathbf{u} \perp \text{span}(W).
\]  

(2.3)

Since \( \text{span}(Z_k) = \text{span}(AQ_k) = \text{span}(MQ_k) \), cf. Eq. (2.2), we choose \( W \) so that \( \text{span}(W) = \text{span}(\nu_0 AQ_k + \mu_0 MQ_k) \), cf. step 5. Here \( \nu_0 \) and \( \mu_0 \) are the weight parameters. As suggested by \([19]\) in step 1 we let \( \nu_0 = 1/\sqrt{1 + |\tau|^2} \), \( \mu_0 = -\tau \nu_0 \), where \( \tau \) is the target. This choice is effective to find the interior eigenvalue close to \( \tau \). The new test vector is first orthogonalized to \( Z \) and then orthogonalized to the current basis of \( W \), cf. step 6.

We note that the subspaces \( \text{span}(V) \) and \( \text{span}(W) \) have the same dimension, say \( j \). The associated \( j \times j \) projected eigenproblem is then

\[
(\eta X^A - \zeta X^M) \mathbf{s} = \mathbf{0}.
\]

(2.4)

Here, matrices \( X^A = W^* A V \) and \( X^M = W^* M V \) are computed in steps 8–11.

By using the QZ decomposition, Eq. (2.4) reduces to generalized Schur form, cf. step 12,

\[
(S^L)^*(X^A)(S^R) = T^A, \quad (S^L)^*(X^M)(S^R) = T^M.
\]

(2.5)

Here, \( S^L \) and \( S^R \) are \( j \times j \) matrices and \( T^A \) and \( T^M \) are \( j \times j \) upper triangular matrices. The decomposition is reordered so that the ratio of \((i,i) (i < j)\) entries of \( T^A \) and \( T^M \), i.e. \( T_{i,i}^A/T_{i,i}^M \), is the \( i \)-th closest to the target \( \tau \). \((T_{i,i}^A, T_{i,i}^M)\) is called the Petrov pair. Then we use \( V S^R \) to approximate \( Q_k \) and \( W S^L \) to approximate \( Z_k \) in Eq. (2.2), respectively. For instance, \( \mathbf{u} = V S^R_1 \) is the approximation of the generalized Schur vector of the first column in \( Q_k \); \( \mathbf{u} \) is called the right Petrov vector and \( \mathbf{p} = W S^L_1 \) is called the left Petrov vector, cf. step 13.

The JDQZ method is capable of computing a number of generalized Schur pairs. Suppose, we have already found \( k-1 \) generalized Schur pairs, i.e. \( AQ_{k-1} = Z_{k-1} R_{k-1}^A, MQ_{k-1} = Z_{k-1} R_{k-1}^M \). If we expand this form with a
2 Electromagnetic eigenvalue problem with loss

Algorithm 1 Jacobi–Davidson QZ algorithm for multiple interior eigenpairs close to a target $\tau$ for GNHEP.

**Input:** Matrices $A, M, Y, H, C, K$; initial vector $v_0$; target $\tau$; tolerance $\varepsilon$; decay parameter $\gamma$; $it_{\text{max}}$, $k_{\text{max}}$, $j_{\text{max}}$, $j_{\text{min}}$.

**Output:** converged eigenpairs $(\lambda_i, q_i)$ ($i = 0, 1, \ldots$, and $i < k_{\text{max}}$) which are closest to $\tau$.

1: Set $t = v_0$; $\nu_0 = 1/\sqrt{1+|\tau|^2}$; $\mu_0 = -\tau\nu_0$; $j = 0$; $T = (A - \tau M)$.
2: Allocate multivectors $Q = [], Z = [], V = [], W = [], V^A = [], V^M = []$.
3: while $it < it_{\text{max}}$ do
4: Project $t$ on $R(C)^\perp$, i.e., $t = (I - YH^{-1}C^T)t$. Then orthogonalyze $t$ against $V$ by using Modified Gram-Schmidt method (MGSM), and normalize $t = t/\|t\|_2$.
5: if $j = j_1 + 1$. $v_j = t$, $v_j^A = Av_j$, $v_j^M = Mv_j$, $w = \mu_0 v_j^A + \mu_0 v_j^M$.
6: Orthogonalize $w$ against $Z$ by using MGSM, then Orthogonalize $w$ against $W$ by using MGSM, and normalize $w = w/\|w\|_2$.
8: for $i = 1, \ldots, j - 1$ do
9: $X^A_{i,j} = w_j^A v_j^A$, $X^M_{i,j} = \delta_j v_i^A$, $X^A_{i,j} = w_j^M v_j^M$, $X^M_{i,j} = \eta_j v_i^M$.
10: end for
11: $X^A_{i,j} = w_j^A v_i^A$, $X^M_{i,j} = \delta_j v_j^A$, $X^A_{i,j} = w_j^M v_j^M$, $X^M_{i,j} = \eta_j v_i^M$.
12: compute the QZ decomposition: $X^A S_R = S_L^T A^M$, $X^M S_M = S^L T M$, such that: $|T_{i,i+1,i+1} - \tau| \leq |T_{i,i+1,i+1}^A - \tau| \leq |T_{i,i+1,i+1}^M - \tau|$, $i = 1, \ldots, j - 1$.
13: $u = Vs^R_i$, $p = Ws^L_i$, $u^A = V^A s^R_i$, $u^M = V^M s^R_i$, $\zeta = T_{i+1,i}^M$, $\eta = T_{i+1,i}^M$.
14: $v = \eta u^A - \zeta u^M$, $\tilde{v} = Z^* u^A$, $\tilde{m} = Z^* u^M$, $\tilde{f} = v - Z(\eta \tilde{v} - \zeta \tilde{m})$.
15: if $\|\tilde{f}\|_2 < \varepsilon$ then
16: $R^A = \begin{bmatrix} R^A & \tilde{a} \\ 0 & \zeta \end{bmatrix}$, $R^M = \begin{bmatrix} R^M & \tilde{m} \\ 0 & \eta \end{bmatrix}$.
18: if $k = k_{\text{max}}$ then break.
19: $j = j_1$.
20: for $i = 1, \ldots, j$ do
21: $u = Vs^R_{i+1}$, $v_i^A = V^A s^R_i$, $v_i^M = V^M s^R_i$, $w_i = Ws^L_{i+1}$, $s_i^R = s_i^L = e_i$.
22: end for
23: $X^A$, $X^M$ is the lower $j \times j$ block of $T^A$, $T^M$, respectively.
24: $u = v_1$, $p = w_1$, $u^A = v_1^A$, $u^M = v_1^M$, $\zeta = T_{1,1}^M$, $\eta = T_{1,1}^M$.
25: $v = \eta u^A - \zeta u^M$, $\tilde{v} = Z^* u^A$, $\tilde{m} = Z^* u^M$, $\tilde{f} = v - Z(\eta \tilde{v} - \zeta \tilde{m})$.
26: end if
27: if $j \geq j_{\text{max}}$ then
28: for $i = 2, \ldots, j_{\text{min}}$ do
29: $v_i = Vs^R_i$, $v_i^A = V^A s^R_i$, $v_i^M = V^M s^R_i$, $w_i = Ws^L_i$.
30: end for
31: $X^A$, $X^M$ is the leading $j_{\text{min}} \times j_{\text{min}}$ block of $T^A$, $T^M$, respectively.
32: $v_1 = u$, $w_1 = p$, $v_1^A = u^A$, $v_1^M = u^M$, $j = j_{\text{min}}$.
33: end if
35: Find an approximate solution of the correction equation
\begin{equation}
(I - \tilde{Z} \tilde{Z}^*) (I - \tilde{Q} \tilde{Q}^*) t = -\tilde{f}, \quad t \perp \tilde{Q}.
\end{equation}
by a Krylov solver preconditioned by $K$.
36: end while
37: Solve $k \times k$ small-sized eigenvalue problem $R^A x_i = \lambda_i R^M x_i$ ($i = 1, \ldots, k$) with the QZ method.
38: Return with the converged eigenpairs $(\lambda_i, q_i = Qx_i)$, $i = 1, \ldots, k$. 

14
2.3 The JDQZ eigensolver

new right vector $q_k$ and a new left vector $z_k$, then

$$A[Q_{k-1}, q_k] = [Z_{k-1}, z_k] \begin{bmatrix} R^A & \tilde{a} \\ 0 & \zeta_k \end{bmatrix},$$

$$M[Q_{k-1}, q_k] = [Z_{k-1}, z_k] \begin{bmatrix} R^M & \tilde{m} \\ 0 & \eta_k \end{bmatrix},$$

(2.7)

$$Q_{k-1}^* q_k = 0, \quad Z_{k-1}^* z_k = 0.$$

Here the two auxiliary vectors are $\tilde{a} = Z_{k-1}^* A q_k$ and $\tilde{m} = Z_{k-1}^* M q_k$. With Eq. (2.7) we have the relation

$$(\eta_k A - \zeta_k M) q_k - Z_{k-1} (\eta_k \tilde{a} - \zeta_k \tilde{m}) = 0.$$

Then, for the Petrov pair $(\eta, \zeta) \approx (\eta_k, \zeta_k)$ and the corresponding Petrov vector $u \approx q_k$, cf. steps 13 and 24, the residual is

$$\tilde{r} = (\eta A - \zeta M) u - Z_{k-1} (\eta \tilde{a} - \zeta \tilde{m}).$$

As long as the norm of $\tilde{r}$ is smaller than the tolerance $\varepsilon$, cf. step 15, we accept the Petrov pair and Petrov vector $((\eta, \zeta), u)$. $Q$ is expanded by the right Petrov vector $u$, and $Z$ is expanded by the associated left Petrov vector $p = W S^L_1$. In order to search for the next, right Schur vector, the search subspace span($V$) can be deflated, such that it is orthogonal to the converged Petrov vector $u$. Therefore, we span the new search subspace using Petrov vectors $VS^R(:, 2), \ldots, VS^R(:, j)$, i.e $V = VS^R(:, 2 : j)$. Due to the change of $V$, multivectors $W, V^A, V^M, S^R, S^L$ and matrices $X^A, X^M$ are all adjusted, cf. steps 20–23.

To improve the performance of the JDQZ algorithm, restart technique is used in order to restrict the dimension $j$ of search subspace span($V$) and test subspace span($W$). As long as $j \leq j_{\text{max}}$, the iteration continues with the two subspaces of smaller dimension. Here, the new search subspace is spanned by the best $j_{\text{min}}$ right Petrov vectors, i.e. $V = VS^R(:, 1 : j_{\text{min}})$, which are associated with Petrov pairs closest to the target value $\tau$. The new test subspace is also spanned by the associated $j_{\text{min}}$ left Petrov vectors, i.e. $W = W S^L(1 : j_{\text{min}})$, cf. steps 28–32. The restart technique restricts the maximum memory consumption. This is essential so that large scale problems can be solved.

A search subspace is expanded by a vector which is the solution of the so-called correction Eq. (2.6). Note that the matrix $\tilde{Q} = [Q, u]$ is $Q$ expanded by the current Petrov vector $u$, cf. step 34. Thus the correction $t$ is required to be orthogonal to the columns of $Q$ as well as to $u$. Eq. (2.6) can be solved
with a preconditioned Krylov subspace method designed for non-Hermitian systems. Here, we implement the biconjugate gradient stabilized (Bi-CGStab) solver \[76\]. The Bi-CGStab solver inherits the advantages of low memory consumption and smart extraction strategy. For any Jacobi-Davidson method the solution of the correction equation is not required to high accuracy; usually, a few iteration steps are sufficient.

Algorithm 2 Solving the correction equation for Jacobi-Davidson QZ method. Approximately solve the correction equation (2.6):
\[(I - \tilde{Z}\tilde{Z}^*)T(I - \tilde{Q}\tilde{Q}^*)t = -\tilde{r}, \quad t \perp \tilde{Q}.
\]

Use the preconditioner (2.8) to approximate \(\tilde{T} = (I - \tilde{Z}\tilde{Z}^*)T(I - \tilde{Q}\tilde{Q}^*)\).

1: solve \(\hat{Z}\) from \(K\hat{Z} = \tilde{Z}\),
2: compute \(X = \tilde{Q}^*\hat{Z}\),
3: LU decomposition: \(X = LU\),
4: compute the preconditioned right-hand side \(\tilde{r} = \tilde{K}^{-1}r\) as follows:
5: (a) solve \(\hat{r}\) from \(K\hat{r} = r\)
6: (b) \(\hat{\gamma} = \tilde{Q}^*\hat{r}\)
7: (c) solve \(\hat{\beta}\) from \(L\hat{\beta} = \hat{\gamma}\)
8: (d) solve \(\hat{\alpha}\) from \(U\hat{\alpha} = \hat{\beta}\)
9: (d) compute \(\tilde{r} = \hat{r} - \hat{Z}\hat{\alpha}\).
10: apply Bi-CGstabil Krylov solver with start vector \(t_0\) \((\tilde{Q}^*t_0 = 0)\). The operator is \(\tilde{K}^{-1}\tilde{T}\), and right-hand side is \(-\tilde{r}\). Compute \(z = \tilde{K}^{-1}\tilde{T}v\) for a given vector \(v\) as follows:
11: (a) \(y = T v\)
12: (a) solve \(\hat{y}\) from \(K\hat{y} = y\)
13: (b) \(\hat{\gamma} = \tilde{Q}^*\hat{y}\)
14: (c) solve \(\hat{\beta}\) from \(L\hat{\beta} = \hat{\gamma}\)
15: (d) solve \(\hat{\alpha}\) from \(U\hat{\alpha} = \hat{\beta}\)
16: (d) compute \(\tilde{z} = \tilde{z} - \tilde{Z}\tilde{\alpha}\).

Alg. 2, taken from \[68\], solves the correction equation (2.6). To improve the convergence speed of the JDQZ solver, we choose the preconditioner for the correction equation with the from
\[\tilde{K} = (I - \tilde{Z}\tilde{Z}^*)K(I - \tilde{Q}\tilde{Q}^*).\] (2.8)

Here, matrix \(K\) approximates \(T\) and is cheap to invert. We start the Bi-CGStab method with the initial guess \(t_0 = 0\) which is obviously orthogonal to \(\tilde{Q}\). All intermediate iteration vectors for the Bi-CGstab solver will be also in
2.4 Complex-valued dielectric material properties

the subspace orthogonal to \( \tilde{Q} \). In that subspace the critical step is to compute the vector \( z = \tilde{K}^{-1} \tilde{T} v \) for a given vector \( v \) (without explicitly computing the inverse of \( \tilde{K} \)), where

\[
\tilde{T} = (I - \tilde{Z} \tilde{Z}^*) T (I - \tilde{Q} \tilde{Q}^*).
\]

This is achieved by steps 11 to 16. Since \( \tilde{Q}^* v = 0 \), we first compute

\[
\tilde{T} v = (I - \tilde{Z} \tilde{Z}^*) T (I - \tilde{Q} \tilde{Q}^*) v = (I - \tilde{Z} \tilde{Z}^*) y,
\]

where \( y = T v \), cf. step 11. Then, we solve \( z \) with preconditioning

\[
\tilde{K} z = (I - \tilde{Z} \tilde{Z}^*) y, \quad z \perp \tilde{Q}.
\]

The condition \( \tilde{Q}^* z = 0 \) leads to \( K z = y - \hat{Z} \bar{\alpha} \), where \( \hat{Z} = K^{-1} \hat{Z} \), cf. step 1. The vector \( \bar{\alpha} \) is

\[
\bar{\alpha} = (\tilde{Q}^* K^{-1} \hat{Z})^{-1} \tilde{Q}^* K^{-1} y = (\tilde{Q}^* \hat{Z})^{-1} \tilde{Q}^* \hat{y} = \chi^{-1} \tilde{\gamma}.
\]

Here, \( \hat{y} = K^{-1} y \), cf. step 12; \( \chi = \tilde{Q}^* \hat{Z} \) (step 2); and \( \tilde{\gamma} = \tilde{Q}^* \hat{y} \), cf. step 13.

Via a similar approach the preconditioned right-hand side \( \tilde{r} = \tilde{K}^{-1} r \) is computed in steps 5–9.

2.4 Complex-valued dielectric material properties

2.4.1 Formulation of the problem

When we employ lossy dielectric materials in resonant cavity structures and discretize the electric curl-curl equation 1.1 with the finite element method, we eventually obtain a generalized non-Hermitian eigenvalue problem (GNHEP)

\[
A x = \lambda M x,
\]

with the constraint

\[
C^T x = 0.
\]

Here, \( A \) is a real-valued symmetric matrix, \( M \) is a complex-valued symmetric matrix [13], and \( C \) is a complex-valued rectangular matrix.

After solving the GNHEP with the JDQZ eigensolver, the electromagnetic fields \( E \) and \( H \) are derived from the calculated eigenvector. The wavenumber in free space \( k_0 \), the corresponding angular frequency \( \omega \) and the decay rate \( \alpha \) are derived from the eigenvalue. Since \( \lambda = k_0^2 \), we have

\[
k_0 = \sqrt{\lambda}, \quad \omega = \text{Re}(k_0) c, \quad \alpha = \text{Im}(k_0) c.
\]
Here, $c$ is the speed of light in free space. In analogy to an oscillating system with damping, the quality factor is defined as \cite{63} \[ Q_1 = \frac{\omega}{2\alpha}. \] (2.12)

The system’s total stored energy $U$ \cite{63} is, under the assumption of non-dispersive dielectric behavior, \[ U = \frac{\varepsilon_0}{2} \int_{\Omega} \text{Re}(\varepsilon_r)|E|^2 \, dx. \] (2.13)

Here, $\Omega$ is the computational domain of the cavity structure. The energy $U_d$ dissipated in lossy dielectrics is directly related to $\text{Im}(\varepsilon_r)$ \cite{63} \[ U_d = \frac{\varepsilon_0}{2} \int_{\Omega} \text{Im}(\varepsilon_r)|E|^2 \, dx. \] (2.14)

We denote by $P_s$ the power dissipated through the current $J_s$ flowing on the conductor’s surface $\Gamma$. In technically relevant situations the surface conductivity $\sigma_s$ approaches large, practically infinite values. Using the perfect electric conductor boundary condition is appropriate. Then, the electric field is orthogonal on the boundary surface $\Gamma$, while the magnetic field only has got a tangential component $H_t$ on $\Gamma$, \cite{63} \[ J_s = H_t \quad \text{A/m}. \] (2.15)

Therefore, $P_s$ is computed by \cite{63} \[ P_s = \frac{1}{2} \int_{\Gamma} R_s |J_s|^2 \, dx = \frac{1}{2} \int_{\Gamma} R_s |H_t|^2 \, dx. \] (2.16)

Here, $R_s$ is the surface resistivity. In this chapter, we use copper for the surface conductor with $\sigma_s = 5.8 \times 10^7$ (S/m). We compute $R_s$ based on the skin-effect model \cite{63}. Later in this chapter, we will show that the dissipated loss on $\Gamma$ is almost negligible when compared to the loss caused by complex-valued dielectrics, i.e. $\omega P_s \ll U_d$. With definitions in Eqs. (2.13), (2.14) and (2.16) the system’s quality factor can also be defined by \[ Q_2 = \frac{U}{U_d + \omega P_s} \approx \frac{U}{U_d}. \] (2.17)

We will show that the two definitions of the quality factor lead to identical results.
2.4 Complex-valued dielectric material properties

2.4.2 Numerical Experiments

All simulations in this section were carried out on the Cray XT6 of the Swiss National Supercomputing Centre (CSCS) [14]. We study an example taken from [13], Fig 2.1. We model a 20 mm $\times$ 20 mm $\times$ 10 mm cavity with a lossy dielectric block 7 mm $\times$ 7 mm $\times$ 8 mm located centrally on the cavity’s square floor. The dielectric constant of the lossy block is $\varepsilon_r = 10 - 2i$.

Figure 2.1: We show the geometry of the lossy dielectric block cavity in (a). We give a cross-sectional view in (b) where the light gray area represents vacuum material properties while the gray area represents the lossy dielectric block.

In Femaxx, both linear and quadratic Nédélec elements are implemented, and both are used for comparison. We employ Jacobi preconditioning, i.e. $K \approx \text{diag}(T)$. We find that the Jacobi preconditioner is efficient enough for this problem.

We study 3 different unstructured tetrahedral meshes, Mesh1, Mesh2 and Mesh3. The number of elements in these meshes is, 7'814, 44'823 and 141'149, respectively. Correspondingly, when linear elements are used, the number of degrees of freedom (dof) for $T$ is 7'770, 46'798 and 153'465, respectively. When quadratic elements are used, the number of dof for $T$ is 44'906, 265'220 and 855'912, respectively. We set $j_{\text{min}} = 10$ and $j_{\text{max}} = 25$, and the tolerance parameter $\varepsilon = 10^{-10}$.
2 Electromagnetic eigenvalue problem with loss

Table 2.1: Numerical results of the lossy dielectric block cavity: linear elements are used. $f$: the resonant frequency $= \omega/(2\pi)$; $Q_1$: the quality factor defined by Eq. (2.12); $Q_2$: the quality factor defined by Eq. (2.17).

<table>
<thead>
<tr>
<th>mode</th>
<th>Mesh1</th>
<th>Mesh2</th>
<th>Mesh3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$f$ (GHz)</td>
<td>$Q_1$</td>
<td>$Q_2$</td>
</tr>
<tr>
<td>1</td>
<td>6.092</td>
<td>10.5</td>
<td>10.5</td>
</tr>
<tr>
<td>2/3</td>
<td>9.109</td>
<td>5.87</td>
<td>5.82</td>
</tr>
<tr>
<td>4/5</td>
<td>11.13</td>
<td>7.99</td>
<td>7.95</td>
</tr>
<tr>
<td>6</td>
<td>11.55</td>
<td>5.54</td>
<td>5.50</td>
</tr>
</tbody>
</table>

Table 2.2: Numerical results of the lossy dielectric block cavity: quadratic elements are used. $f$: the resonant frequency $= \omega/(2\pi)$; $Q_1$: the quality factor defined by Eq. (2.12); $Q_2$: the quality factor defined by Eq. (2.17).

<table>
<thead>
<tr>
<th>mode</th>
<th>Mesh1</th>
<th>Mesh2</th>
<th>Mesh3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$f$ (GHz)</td>
<td>$Q_1$</td>
<td>$Q_2$</td>
</tr>
<tr>
<td>1</td>
<td>6.160</td>
<td>11.1</td>
<td>11.0</td>
</tr>
<tr>
<td>2/3</td>
<td>9.093</td>
<td>5.82</td>
<td>5.78</td>
</tr>
<tr>
<td>4/5</td>
<td>11.40</td>
<td>7.49</td>
<td>7.46</td>
</tr>
<tr>
<td>6</td>
<td>11.42</td>
<td>5.49</td>
<td>5.44</td>
</tr>
</tbody>
</table>

The numerical results for the six lowest-order modes are given in Tab. 2.1 and 2.2. Several degenerate modes exist: e.g. modes 2 and 3 are degenerated and they nearly have identical resonance frequencies and quality factors. Our eigenmodal solutions agree well for the resonance frequencies given in [13]; no quality factors are given in that publication.

We note that the quality factors $Q_1$ and $Q_2$, defined in two different ways, show excellent agreement. The energy dissipated on the conductor surface, $\omega P_s$, is smaller than 0.1% of the energy $U_d$ dissipated inside the lossy dielectric block.

The numerical results for the quadratic elements on the Mesh3 are the most accurate, such that we use them as the reference solutions. Then, the quadratic elements on the coarse Mesh1 are almost as accurate, below 0.2% for $f$ and 1% for $Q_1$, when compared to the linear elements on the fine Mesh3. The results with the linear elements on Mesh1 differ from the reference solutions by up to about 2.5% for $f$ and 7% for $Q_1$. In conclusion, quadratic elements are more accurate than linear elements. Therefore we will use quadratic elements.
in the rest of this thesis unless otherwise specified. The electromagnetic field distributions are shown in Figure 2.2.

In table 2.3, we employ quadratic elements on Mesh3 and study the parallel efficiency $E(p)$ of the JDQZ eigensolver where $t = t(p)$ denotes the execution time to solve the eigenvalue problem with a varying number $p$ of compute cores. These times do not include preparatory work, such as the assembly of matrices or data redistribution. For each $p$, the numerical results are exactly the same when compared to the reference solutions. Approximately 135 JDQZ outer iteration steps are required to compute all 6 eigenmodes.

Table 2.3: We show the parallel efficiency $E(p)$ of the JDQZ eigensolver. Experiments are carried out for quadratic elements on Mesh3 ($\#\text{dof} = 855'912$)

<table>
<thead>
<tr>
<th>p</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$ (sec.)</td>
<td>2710</td>
<td>1252</td>
<td>693</td>
<td>373</td>
<td>296</td>
<td>209</td>
<td>187</td>
</tr>
<tr>
<td>$E(p)$</td>
<td>1.00</td>
<td>1.08</td>
<td>0.98</td>
<td>0.91</td>
<td>0.57</td>
<td>0.41</td>
<td>0.23</td>
</tr>
</tbody>
</table>

When using 64 cores or less, the parallel efficiency is above 90%. Particularly, the so-called super linear speedup occurs when $p = 16$, with associated efficiency $> 1$. However, when $p \geq 128$ cores, the communication overhead among different cores increases significantly. Consequently, $E(p)$ drops drastically.

In this situation, it is advantageous to use 64 cores so that $E(64)$ is high while the total execution time is approximately 6 minutes; that is $855'912/64 \approx 13'374$ dof per core. For all remaining simulations in the rest of this thesis, we employ a number of cores so that the number of dof per core is around $10'000-15'000$. This strategy is found to be efficient for using Femaxx. It only requires moderate computational resources, while limiting the total execution time.

2.5 Ohmic conductivity loss in volume regions

2.5.1 Formulation of the quadratic eigenvalue problem

Here, we study electromagnetic loss in resonant systems caused by volume ohmic conductivity $\sigma$. In the low-frequency microwave range, it is justified to exclusively consider real-valued $\sigma$ [63]. We comment that $\sigma$ is a complex
Figure 2.2: We show the electromagnetic field distribution of the cavity which contains a lossy dielectric block; in particular, we show the two lowest-order modes, visualized in the plane where the lossy dielectric block is attached to the metallic wall, cf. Fig. 2.1.
2.5 Ohmic conductivity loss in volume regions

value in general. In particular, if a structure’s resonance resides in the high-frequency, optical range of the electromagnetic spectrum, the imaginary part of $\sigma$ must not be ignored. We derive the electric field vector wave equation considering loss caused by volume ohmic conductivity. We start with Ampere’s equation including the ohmic current density $J$.

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \quad \mathbf{J} = \sigma \mathbf{E}.$$  

Here, for the electric displacement field $\mathbf{D}$, we assume linear, homogeneous and isotropic material properties, so that $\mathbf{D} = \varepsilon_0 \varepsilon_r \mathbf{E}$. Differing from Eq. (1.1a) for lossless systems, the time-harmonic electric field curl-curl equation then becomes

$$\nabla \times (\mu_r^{-1} \nabla \times \mathbf{E}(\mathbf{x})) + ik_0 \sigma Z_0 \mathbf{E}(\mathbf{x}) - k_0^2 \varepsilon_r \mathbf{E}(\mathbf{x}) = 0,$$

$$\nabla \cdot (\varepsilon_r \mathbf{E}(\mathbf{x})) = 0,$$

$$\mathbf{x} \in \Omega.$$  

Here, $\Omega \subset \mathbb{R}^3$ is a bounded domain, $k_0(= \tilde{\omega} \sqrt{\mu_0 \varepsilon_0})$ is the complex wavenumber in free space; $\tilde{\omega}(= \omega + i\alpha)$ is the complex angular frequency, $\omega$ is the angular frequency and $\alpha$ is the exponential decay rate, Eq. (2.11). There appears a term associated with $\sigma$ arising in Eq. (2.18) and $Z_0(= \sqrt{\mu_0 / \varepsilon_0} \approx 377 \Omega)$ is the characteristic impedance of free space [63].

The divergence-free constraint Eq. (1.1b) still holds, and again the perfect electric conductor boundary condition (1.1c) are used. Then, a natural weak form [85] for Eqs. (2.18), (1.1b) and (1.1c) is

$$\int_\Omega \mu_r^{-1} \nabla \cdot \mathbf{f} \cdot \nabla \times \mathbf{E} d\mathbf{x} + ik_0 \int_\Omega \sigma Z_0 \mathbf{f} \cdot \mathbf{E} d\mathbf{x} + k_0^2 \int_\Omega \varepsilon_r \mathbf{f} \cdot \mathbf{E} d\mathbf{x} = 0,$$

$$\int_\Omega \varepsilon_r \mathbf{E} \cdot \nabla q d\mathbf{x} = 0.$$  

Here, $V$ denotes the functions in $H(\text{curl}; \Omega)$ that satisfy the boundary condition (1.1c) and $W = H_0^1(\Omega)$; for details see, e.g., [51] or [39].

Again, we employ Nédélec elements to approximate $V$ and Lagrange elements to approximate $W$. Eq. (2.19) then yields a constrained complex quadratic eigenvalue problem

$$T(\lambda)\mathbf{x} = A\mathbf{x} + \lambda R\mathbf{x} - \lambda^2 M\mathbf{x} = 0,$$

$$C^T \mathbf{x} = 0,$$  

(2.20)
Electromagnetic eigenvalue problem with loss

where \( \lambda (= k_0) \) is the eigenvalue, and \( x \) is the eigenvector. The entries of matrices \( A, M \) and \( C \) are the same as in Eq. (1.5). The element of the newly arising matrix \( R \) is given in Eq. (2.21). We note that \( R \) is a purely imaginary valued matrix.

\[
r_{ij} = i \int_{\Omega} \sigma Z_0 \mathbf{N}_i \cdot \mathbf{N}_j \, d\mathbf{x}, \quad 1 \leq i, j \leq n. \tag{2.21}
\]

Once Eq. (2.20) has been solved, the quality factors are again computed via Eqs. (2.12) or (2.17). The dissipated energy \( U_d \), in Eq. (2.14), is now expressed via \( \sigma \) instead of \( \text{Im}(\varepsilon_r) \) \[63\]

\[
U_d = \frac{1}{2} \int_{\Omega} \mathbf{J} \cdot \mathbf{E} \, d\mathbf{x} = \frac{1}{2} \int_{\Omega} \sigma |\mathbf{E}|^2 \, d\mathbf{x}. \tag{2.22}
\]

2.5.2 Linearization of the quadratic eigenvalue problem (QEP)

We solve the quadratic eigenvalue problem (2.20) through an appropriate linearization \[74\]. Defining \( x_2 = \lambda x_1 \), Eq. (2.20) is transformed into a constrained linear eigenvalue problem of the form

\[
\mathcal{A}x \equiv \begin{bmatrix} A & O \\ O & I \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} -R & M \\ I & O \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \equiv \lambda \mathcal{B}x, \tag{2.23}
\]

\[
\mathcal{C}^T x \equiv \begin{bmatrix} C & O \\ O & C \end{bmatrix}^T x = 0.
\]

The identity matrix \( I \) in (2.23) could be replaced by any nonsingular matrix \( X \), i.e., we could write \( Xx_2 = \lambda Xx_1 \). Choosing \( X = M \) would make both \( \mathcal{A} \) and \( \mathcal{M} \) complex symmetric. Since complex symmetry does not necessarily increase numerical simplicity \[5\] we set \( X = I \).

The linearization of the quadratic eigenvalue problem allows us to use the JDQZ eigensolver. Now, the global matrix of the correction equation (2.6) has a \( 2 \times 2 \) block structure. The critical part of the correction equation is solving for \( z \) in

\[
Tz = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad \text{where } T = \mathcal{A} - \tau \mathcal{B} = \begin{bmatrix} A + \tau R & -\tau M \\ -\tau I & I \end{bmatrix}. \tag{2.24}
\]

Here, \( \tau \) is the eigensolver target. Instead of solving with \( T \), we carry out one block Gauss–Seidel iteration step,

\[
\begin{align*}
z_1 &= T_{11}^{-1} f_1, \\
z_2 &= T_{22}^{-1} (f_2 - T_{21} z_1).
\end{align*}
\]

24
2.5 Ohmic conductivity loss in volume regions

To solve for $z_1$, we approximate the block $T_{11}$ by its diagonal, i.e., Jacobi preconditioning. Solving for $z_2$ is trivial. We note that $T_{22} = I$ and $T_{21} = -\tau I$, so

$$z_2 = f_2 + \tau z_1.$$ 

### 2.5.3 Numerical experiments

We study two different resonator structures with ohmic conductivity loss. Both examples were originally analyzed in [85]. In both cases we use perfect electric conductor (PEC) boundary conditions. We compute the power dissipated on the conductor surface via the a posteriori approximation that uses the tangential magnetic field to define the surface current density, Eq. (2.16).

#### Half-filled rectangular resonator

We analyze a rectangular cavity resonator [85], where one half of the complete volume is filled with ohmically lossy material and the other half corresponds to vacuum material properties. The dimensions of the complete cavity are 22.86, 22.86, 10.16 mm, respectively. Half of the resonator volume, $11.43 \times 22.86 \times 10.16 \text{ mm}^3$, is filled with ohmically lossy material, where the dielectric permittivity is $\varepsilon_r = 2.0$ and the ohmic conductivity $\sigma$ assumes different values, Tab. 2.4. The resonator’s walls are made from copper.

We generate a mesh using Gmsh [24] with 306'001 tetrahedra. When using quadratic elements, the dof count is 1'852'810 for $A$, $R$ or $M$, and the dof count for $T$ in Eq. (2.24) is twice as large, namely 3'705'620. On 256 cores, to compute the dominant mode requires less than 3 minutes with a residual smaller than $10^{-7}$. The numerical results are listed in Tab. 2.4.

Table 2.4: The numerical results the dominant mode of the half-filled rectangular resonator. $\tilde{\omega}$: the complex angular frequency (Unit: GHz), $\tilde{\omega} = \omega + i\alpha$; $Q_1$: the quality factor defined by Eq. (2.12); $Q_2$: the quality factor defined by Eq. (2.17).

<table>
<thead>
<tr>
<th>$\sigma$ (S/m)</th>
<th>analytical solution for $\tilde{\omega}/(2\pi)$</th>
<th>$\tilde{\omega}/(2\pi)$ computed by Femaxx</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>$7.379 + j0.354$</td>
<td>$7.379 + j0.354$</td>
<td>10.43</td>
<td>10.39</td>
</tr>
<tr>
<td>0.5</td>
<td>$7.236 + j1.819$</td>
<td>$7.236 + j1.817$</td>
<td>1.992</td>
<td>1.989</td>
</tr>
<tr>
<td>1.0</td>
<td>$6.579 + j3.864$</td>
<td>$6.579 + j3.864$</td>
<td>0.5494</td>
<td>0.5493</td>
</tr>
<tr>
<td>1.3</td>
<td>$5.711 + j5.197$</td>
<td>$5.711 + j5.198$</td>
<td>0.8514</td>
<td>0.8511</td>
</tr>
</tbody>
</table>

We compare the numerical values for $\tilde{\omega}$ with the analytical solution from [85]. The relative differences are almost 0. The quality factors $Q_1$ and $Q_2$, again, show excellent agreement. When $\sigma$ increases, the dissipative energy loss $U_d$
2 Electromagnetic eigenvalue problem with loss

also increases, thus the quality factor is reduced. The electromagnetic fields of the dominant mode are shown in Fig. 2.3.

![Electromagnetic field distributions](image)

(a) $|E|$  
(b) $|H|$

Figure 2.3: We show the electromagnetic field distribution of the dominant mode in a rectangular resonator whose right half is filled with ohmically lossy material: $\varepsilon_r = 2.0$ and $\sigma = 1.3$ (S/m).

Cylindrical cavity containing a conductive rod with circular cross-section

Here, we study a cylindrical cavity resonator with circular cross-section which in turn contains an ohmically conducting rod also with circular cross-section. The material properties of the conducting rod are $\varepsilon_r = 37.6$ and $\sigma = 0.1$ S/m [85], respectively. The radius of the cavity resonator is $R = 12.7$ mm, its length is $l = 13.97$ mm. The radius of the conducting rod is $r = 10.0076$ mm. We show the geometry in Fig. 2.4. The cavity walls are modeled with copper.

Using Gmsh [24] we generate a tetrahedral mesh with 156'895 elements. When using quadratic elements, the dof count is 947'252 for $A$, $R$ or $M$, and it is twice as large for $T$, namely 1'894'504. Using 128 cores the time to compute the five lowest-order eigenmodes is approximately 20 minutes, with residuals smaller than $3 \times 10^{-6}$.

The numerical results are shown in Tab. 2.5. Modes 2 and 3 are degenerate, so are modes 4 and 5. Results in $\tilde{\omega}/(2\pi)$ show excellent agreement with a previous numerical solution in [85], as do the quality factors $Q_1$ and $Q_2$. With every mode, the energy dissipated on the surface conductor, $\omega P_s$, is below 1% of the energy dissipated inside the rod, $U_d$. The electric field distributions of modes 1 and 4 are shown in Figs. 2.5(a) and 2.5(b), respectively.
2.5 Ohmic conductivity loss in volume regions

Figure 2.4: We show the geometry of the cylindrical cavity with circular cross-section containing in turn an ohmically conducting rod, marked by the red area.

Table 2.5: We show the numerical results of the five lowest-order eigenmodes for the cylindrical cavity containing a conductive rod with circular cross-section. $\tilde{\omega}$: the complex angular frequency (Unit: MHz), $\tilde{\omega} = \omega + i\alpha$; $Q_1$: the quality factor defined by Eq. (2.12); $Q_2$: the quality factor defined by Eq. (2.17).

<table>
<thead>
<tr>
<th>mode</th>
<th>numerical solution: $\tilde{\omega}/(2\pi)$ in [86]</th>
<th>$\tilde{\omega}/(2\pi)$ computed by Femaxx</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1497.3 + j23.88</td>
<td>1497.3 + j23.89</td>
<td>31.3</td>
<td>31.1</td>
</tr>
<tr>
<td>2</td>
<td>2433.9 + j23.88</td>
<td>2433.6 + j23.94</td>
<td>50.8</td>
<td>50.6</td>
</tr>
<tr>
<td>3</td>
<td>2433.9 + j23.87</td>
<td>2433.5 + j23.89</td>
<td>51.0</td>
<td>50.6</td>
</tr>
<tr>
<td>4</td>
<td>2502.2 + j23.71</td>
<td>2502.0 + j23.70</td>
<td>52.8</td>
<td>52.6</td>
</tr>
<tr>
<td>5</td>
<td>2502.2 + j23.74</td>
<td>2502.2 + j23.62</td>
<td>53.0</td>
<td>52.6</td>
</tr>
</tbody>
</table>
2 Electromagnetic eigenvalue problem with loss

Figure 2.5: We show the electric field distributions for the cylindrical cavity containing a conductive rod with circular cross-section: (a) mode 1 and (b) mode 4.

2.6 Summary

We have developed a parallel Jacobi-Davidson QZ (JDQZ) eigensolver for solving the generalized non-Hermitian eigenvalue problem. We have solved the quadratic eigenvalue problem by linearization. In order to demonstrate the algorithm’s capability, we analyze two different types of resonating cavities that contain volume blocks of lossy dielectric and ohmically conducting materials, respectively. We present the numerical results, complex resonance frequency and quality factor, and compare them to analytical and numerical reference solutions, respectively. In either case we have obtained excellent agreement. We also show the electromagnetic field distributions of some lowest-order modes. We comment that Femaxx also encounters degenerate modes and handles them correctly.
Plasmonic nanostructures

The material presented in this chapter has been published as a paper in the peer-reviewed journal *Optics Express* [29]. Here, we introduce a 3-dimensional electromagnetic eigenmodal algorithm for the theoretical analysis of resonating nano-optical structures. The method, a variant of the Jacobi–Davidson algorithm, solves the electric field vector wave, or curl-curl, equation for the electromagnetic eigenmodes of resonant optical structures with a finite element method. In particular, the method includes transparent boundary conditions that enable the analysis of resonating structures in unbounded space. We demonstrate the performance of the method. First, we calculate the modes of several dielectric resonator antennas and compare them to theoretically determined results. Second, we calculate the modes of a nano-cuboid and compare them to theoretically determined results. Third, we numerically analyze spherical nanoparticles and compare the result to the theoretical Mie solution. Fourth, we analyze optical dipole antenna configurations in order to assess the method’s capability for solving technologically relevant problems.

### 3.1 Introduction

The interaction of nano-meter structured, metallic, aka. plasmonic, devices with light, especially from lasers, has attracted considerable interest during the last few years. Consequentially, a significant body of literature exists, see e.g. [54, 53, 9, 7]. The concept of the antenna, in particular, transferred from the microwave to the optical region, has attracted enormous interest since the conversion of propagating light into localized, electromagnetic energy, and vice-versa, holds the promise of novel technological applications in many fields, such as in sensing, detection manipulation and communication [2, 62, 44, 22]. Certainly, it is well beyond the scope of this study to give a comprehensive overview. Simultaneously, fabrication technology has advanced impressively and is now capable of producing ultra-smooth structures with justifiable effort [52].
3 Plasmonic nanostructures

Since the behavior of nano-optical devices is sometimes counter-intuitive and the fabrication processes are often time-consuming, the theoretical analysis of these devices has become indispensable in order to identify promising candidates and to assess their performance before resources are committed in the clean room.

Generally speaking, theoretical methods are classified into analytical [33], semi-analytical [31] and numerical approaches [39, 51, 79]. While analytical methods can provide detailed insight into the behavior of devices and allow for relatively simple parameter sweeps, quite often their applicability is limited to simplified geometries and, thus, their predictions are of limited validity when technologically relevant problems are studied. Numerical methods on the other hand allow for the detailed modeling of delicate geometrical details, almost fully preserving the reality of the device. Thus, they are the methods of choice for detailed performance analysis. We note however that such detailed wealth of prediction comes at a price in the form of significantly increased computational effort and substantial pre- and post-processing work. In between these methods is the class of semi-analytical methods that are usually more flexible with respect to geometry, but on the other hand still provide analytical insight. Among the many semi-analytical methods we particularly mention the multiple multipole (MMP) method, originated by Hafner [31] in the 1980s. The MMP method subdivides geometry into subdomains for which the analytical solutions are available. It then matches the superposition of these scaled analytical solutions on the interfaces between the subdomains by formulating the mismatch as an error energy. The resulting, overdetermined linear system is then solved and the mismatch on the subdomain interfaces is a direct measure for accuracy.

Here, we use the finite element method (FEM) [39] in the time-harmonic regime to calculate electromagnetic modes associated with resonant nano-optical devices. We derive and solve the corresponding electromagnetic eigenmodal problem. The resonance wavelength, the decay rate and the quality factor (Q factor) are derived from the (complex) eigenvalue, and the electromagnetic fields are derived from the calculated eigenvector. The radiated power, dissipated energy and stored energy of the system can be computed as well. We demonstrate the performance of the method by solving four different problems, namely:

(i) The dielectric resonator antenna (DRA) [57, 50]: the problem is solved in the microwave region. A theoretical model, applicable under specific conditions, is available to benchmark the method.

(ii) The single nano-cuboid: the problem is solved in the optical region. We compare the numerical solution with the theoretical model proposed in
3.2 Formulation of the problem

(iii) The single spherical nanoparticle: we compare the numerical solution with the Mie solution [10].

(iv) The optical dipole antenna of finite lateral dimensions: we realistically model the dipole antenna with two smoothly rounded arms. This geometry has traditionally been studied via the thin wire approximation [63]. The FDTD method [16] experiences intrinsic difficulties when modeling rounded objects. Other work [15, 48, 70] is restricted to 2-dimensional geometries. In [71, 43], the 3-D FEM is restricted to relatively small-scale problems due to the huge memory consumption and excessive computation time, further suffering from relatively low accuracy. In [41, 42], the surface integral equation (SIE) method experiences difficulties when modeling the substrate. Here, we study a technologically relevant geometrical discretization of the optical dipole antenna, free of the aforementioned limitations. Both bright and dark modes [12] are investigated.

3.2 Formulation of the problem

The electromagnetic background of our simulations originates in the Maxwell equations combined with proper boundary conditions. In the time-harmonic regime, after eliminating the magnetic field \( \mathbf{H}(x) \), the electric field \( \mathbf{E}(x) \) satisfies

\[
\nabla \times (\mu_r^{-1} \nabla \times \mathbf{E}(x)) + ik_0 \sigma Z_0 \mathbf{E}(x) - k_0^2 \varepsilon_r \mathbf{E}(x) = 0, \quad \nabla \cdot (\varepsilon_r \mathbf{E}(x)) = 0, \quad x \in \Omega.
\]

Here, \( \Omega \subset \mathbb{R}^3 \) is a bounded domain, \( k_0(= \tilde{\omega} \sqrt{\mu_0 \varepsilon_0}) \) is the complex wavenumber in free space; \( \tilde{\omega}(= \omega + i\alpha) \) is the complex angular frequency, \( \omega \) is the angular frequency and \( \alpha \) is the exponential decay rate; \( \mu_r, \varepsilon_r \) are relative magnetic permeability and relative electric permittivity, respectively; \( \sigma \) is the ohmic loss of the material; \( Z_0(= \sqrt{\mu_0 / \varepsilon_0}) \) is the characteristic impedance of free space.

We model the material loss of the nano metallic structure through the complex permittivity \( \varepsilon_r \), normally with negative real part, instead of ohmic loss \( \sigma \). The dielectric function is not generally constant, but \( \varepsilon_r \) depends on the frequency \( \omega \), or, equivalently, the real part of \( k_0 \) [40]. In this paper, we use experimental data from Johnson and Christy [40] and employ a piecewise linear interpolation. We obtain the dielectric function within the spectral range from 0.64 eV to 6.60 eV [40]. The substrate’s, e.g., glass, or the environment’s, e.g. a covering layer of water, dielectric functions are not constant either [32]. However, since the variation is usually small within a certain optical spec-
In the far range, we assume that the dielectric properties of the substrate and the environment are constant.

![Figure 3.1: A 3-D view of a typical computational domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$. $\Omega_1$ is the optical device, $\Omega_2$ is the substrate, $\Omega_3$ is the environment. $\Gamma_{13}$ is the interface between $\Omega_1$ and $\Omega_2 \cup \Omega_3$, and $\Gamma_{23}$ is the interface between $\Omega_2$ and $\Omega_1 \cup \Omega_3$. $\Gamma$ is the boundary of $\Omega$.](image)

The formulation of the nano-optical device simulation uses a bounded domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$. It encloses the nano-optical device $\Omega_1$, its substrate $\Omega_2$, and the environment $\Omega_3$, cf. Fig. 3.1. We solve the eigenproblem Eq. (3.1), where

$$
\varepsilon_x = \begin{cases} 
\varepsilon_r(\text{Re}(k_0)), & \mathbf{x} \in \Omega_1; \\
\varepsilon_{\text{sub}}, & \mathbf{x} \in \Omega_2; \\
\varepsilon_{\text{env}}, & \mathbf{x} \in \Omega_3.
\end{cases} \tag{3.2}
$$

$\varepsilon_{\text{sub}}$ and $\varepsilon_{\text{env}}$ are the constant relative permittivity for the substrate and the environment, respectively. $\varepsilon_r$ depends on $\text{Re}(k_0)$ in the subdomain $\Omega_1$.

On the interface $\Gamma_{13}$ the tangential component of $\mathbf{E}(\mathbf{x})$ must be continuous. On the artificial surface $\Gamma$ we use the 1st order absorbing, aka. transparent, boundary condition (ABC) [39]

$$
\mathbf{n} \times \nabla \times \mathbf{E}(\mathbf{x}) = -i k \mathbf{n} \times (\mathbf{n} \times \mathbf{E}(\mathbf{x})), \tag{3.3}
$$

where $k = k_0 \sqrt{\mu_r \varepsilon_r}$ is the wavenumber of the surrounding medium. We let $\Omega$ be a sphere of radius $R_b$. $R_b$ should be as large as possible to minimize the artificial reflections on $\Gamma$ introduced by the ABC, which arise in the case of non-orthogonal incidence of out-going wave. On the other hand, $R_b$ should be small to limit the size of the discretized system. By numerical experiment we
3.2 Formulation of the problem

fix $R_b$ such that a change of $R_b$ does not result in a significant deviation of the computed resonant wavelength, cf. the discussion in Sections 4.3. Following this rule, we observe that the minimum distance between $\Gamma$ and the device $\Omega_1$ is at least (and usually larger than) $1/3$ of the resonance wavelength.

We solve (3.1)–(3.3) to determine the electric field $E(x)$ and the wavenumber $k_0$. The corresponding magnetic field is obtained by

$$H(x) = \frac{1}{-i\tilde{\omega}\mu_0\mu_r} \nabla \times E(x) = \frac{i}{k_0Z_0\mu_r} \nabla \times E(x).$$

(3.4)

The resonant frequency is $f = \omega/2\pi$ ($\omega = \text{Re}(\tilde{\omega})$). The decay rate $\alpha (= \text{Im}(\tilde{\omega}))$ is directly connected to the loss. In analogy to an oscillating system with damping, the quality factor is defined as

$$Q_1 = |\omega/2\alpha|.$$

(3.5)

We compute the system’s total energy $U$, stored energy $U_s$, dissipated energy $U_d$, and the radiated power $P_r$ of the dispersive metallic nano-structure, respectively. Under the condition of no magnetic dispersion ($\mu_r = 1$), the total energy is [64, 46]

$$U = \frac{\varepsilon_0}{4} \int_\Omega (\text{Re}(\varepsilon_r) + \frac{2\omega}{\gamma_e} \text{Im}(\varepsilon_r)) |E(x)|^2 \, dx + \frac{\mu_0}{4} \int_\Omega |H(x)|^2 \, dx.$$ 

(3.6)

where $\gamma_e$ is the damping frequency of the dispersive metal with $\gamma_e = \tau^{-1}$ where $\tau$ is the relaxation time. $\tau \approx 31$ fs for silver and $\tau \approx 9.3$ fs for gold [40].

If (i) the metal dielectric function can be described by the Drude model [10], and (ii) $\omega\tau \gg 1$, we obtain the stored energy [56, 61] in (3.7), equivalent to the Brillouin formula [11],

$$U_s = \frac{\varepsilon_0}{4} \int_\Omega (\text{Re}(\varepsilon_r) + \omega \frac{d(\text{Re}(\varepsilon_r))}{d\omega}) |E(x)|^2 \, dx + \frac{\mu_0}{4} \int_\Omega |H(x)|^2 \, dx.$$ 

(3.7)

We note that computing the stored energy for nano-metallic structures is still a matter of intense, scientific debate [61, 56]. We emphasize that Eq. (3.7) holds only if both assumptions (i) and (ii) are satisfied. The second assumption, i.e., $\omega\tau \gg 1$, holds for all eigenmodes computed in this paper, e.g., $\omega\tau > 100$ for silver spheres, Section 4.3, and $\omega\tau > 20$ for gold dipole antennas, Section 4.4. While the Drude model is a good fit for dispersive silver, still, there is a difference between the analytical model and the experimental data from [40], particularly noticeable in the range of 2.5 eV to 3.5 eV [56], cf. Section 4.3. For the optical dispersion of gold, the Drude model is a rough fit. In particular, the difference between the analytical model’s imaginary part
of $\varepsilon_r$, and the experimental data is considerable for the photon energy above 2 eV [78]. Therefore, in the spectral range above 2 eV the calculation of the stored energy via Eq. (3.7) is invalid for gold, cf. discussion in Section 4.4.2. However, a study on how to accurately evaluate the stored energy for metals is definitely beyond the scope of this paper. Nevertheless, we employ (3.7) in order to assess its accuracy by comparing it to results derived via an alternative route. Dissipated energy [56] is then computed as

$$U_d = \frac{\varepsilon_0}{2} \int_{\Omega_1} \text{Im}(\varepsilon_r)|\mathbf{E}(\mathbf{x})|^2 \, d\mathbf{x}. \quad (3.8)$$

Radiated power is calculated via the Poynting theorem [63],

$$P_r = \frac{1}{2} \int_{\Gamma} \text{Re}(\mathbf{E}(\mathbf{x}) \times \mathbf{H}(\mathbf{x})^*) \cdot d\mathbf{x}. \quad (3.9)$$

Therefore the system’s quality factor can also be defined by

$$Q_2 = \frac{\omega U_s}{\omega U_d + P_r}. \quad (3.10)$$

We comment that $Q_2$ will appear to be a less reliable definition than $Q_1$, since $U_s$ in (3.7) may not be accurate enough for nano-metallic structures. Therefore, in our numerical examples, we list both $Q_1$ and $Q_2$ for comparison purposes. The radiative quantum yield $\eta$ is evaluated as the ratio of radiation loss and total loss, i.e.,

$$\eta = \frac{P_r}{P_r + \omega U_d}. \quad (3.11)$$

In the special case of a non-dispersive medium that has no dissipative loss, e.g., the dielectric resonator antenna (DRA) with constant $\text{Re}(\varepsilon_r)$ and $\text{Im}(\varepsilon_r) = 0$, (3.7) and (3.8) lead to

$$U_s = \frac{\varepsilon_0}{2} \int_{\Omega} \text{Re}(\varepsilon_r)|\mathbf{E}(\mathbf{x})|^2 \, d\mathbf{x}, \quad U_d = 0, \quad (3.12)$$

such that $Q_2 = \omega U_s / P_r$ and $Q_1 = Q_2$.

### 3.3 Numerical method

#### 3.3.1 The finite element method

The finite element method (FEM) is a suitable method for the nano-optical devices simulation, since such devices exhibit delicate geometries with a wide
span of geometrical scales. In order to apply the finite element method we need a weak form of (3.1). We first note that for sufficiently smooth vector functions $\mathbf{E}(\mathbf{x})$ and $\mathbf{f}(\mathbf{x})$ that satisfy the boundary condition (3.3) we have [39, 63]

$$\int_{\Omega} \mathbf{f}(\mathbf{x}) \cdot (\nabla \times \mu^{-1}_r \nabla \times \mathbf{E}(\mathbf{x})) \, d\mathbf{x}$$

$$= \int_{\Omega} \mu^{-1}_r \nabla \times \mathbf{f} \cdot \nabla \times \mathbf{E} \, d\mathbf{x} - \int_{\Gamma} \mu^{-1}_r \mathbf{n} \cdot (\mathbf{f} \times \nabla \times \mathbf{E}) \, d\mathbf{x}$$

$$= \int_{\Omega} \mu^{-1}_r \nabla \times \mathbf{f} \cdot \nabla \times \mathbf{E} \, d\mathbf{x} + i \int_{\Gamma} k \mu^{-1}_r (\mathbf{n} \times \mathbf{f}) \cdot (\mathbf{n} \times \mathbf{E}) \, d\mathbf{x}.$$

Similarly, for any sufficiently smooth scalar function $q(\mathbf{x})$ vanishing on $\Gamma$ we get

$$\int_{\Omega} \varepsilon_r (\nabla \cdot \mathbf{E}(\mathbf{x})) q(\mathbf{x}) \, d\mathbf{x} = - \int_{\Omega} \varepsilon_r \mathbf{E}(\mathbf{x}) \cdot (\nabla q(\mathbf{x})) \, d\mathbf{x} = 0.$$

Therefore, a natural weak form of (3.1)–(3.3) is:

Find $k_0 \in \mathbb{C}$ and $\mathbf{E} \in V$, $\mathbf{E} \neq \mathbf{0}$, such that for all $\mathbf{f} \in V$ and all $q \in W$

$$\int_{\Omega} \mu^{-1}_r \nabla \times \mathbf{f} \cdot \nabla \times \mathbf{E} \, d\mathbf{x} + ik_0 \left[ \int_{\Gamma} \sqrt{\varepsilon_r} \mu_r (\mathbf{n} \times \mathbf{f}) \cdot (\mathbf{n} \times \mathbf{E}) \, d\mathbf{x} + \int_{\Omega} \sigma Z_0 \mathbf{f} \cdot \mathbf{E} \, d\mathbf{x} \right]$$

$$+ k_0^2 \int_{\Omega} \varepsilon_r \mathbf{f} \cdot \mathbf{E} \, d\mathbf{x} = 0,$$  \hspace{1cm} (3.13)

$$\int_{\Omega} \varepsilon_r \mathbf{E} \cdot \nabla q \, d\mathbf{x} = 0.$$  \hspace{1cm} (3.14)

Here, $V$ denotes the functions in $H(\text{curl}; \Omega)$ that satisfy the boundary condition (3.3) and $W = H^1_0(\Omega)$. For details on these function spaces see, e.g., [51] or [39].

We discretize (3.13) by the Ritz-Galerkin method [39, 51] employing appropriate finite element subspaces of $V$ and $W$. To that end we triangulate $\Omega$ by tetrahedra with Gmsh [24]. The triangulation has to respect the interfaces $\Gamma_{13}$ and $\Gamma_{23}$. The mesh must be fine in the vicinity of the surface of the optical device $\Omega_1$, but it can be coarse in the far-field zone. This strategy maintains efficiency and accuracy while, on the other hand, it reduces the computational cost.

The electric/magnetic vector functions in $V$ are then approximated by Nédelec (edge) elements, while the scalar functions in $W$ are approximated by Lagrange (nodal) finite elements [51]. This approach avoids the generation of spurious eigensolutions, and imposing the boundary conditions is straightforward [39].
3 Plasmonic nanostructures

Let the vector functions \( N_i, 1 \leq i \leq n \), be the Nédélec basis functions, while the scalar functions \( N_{\ell}, 1 \leq \ell \leq m \), denote the Lagrange basis functions. Then, equations (3.13) yield a constrained complex quadratic eigenvalue problem

\[
T(\lambda)x = Ax + \lambda R x - \lambda^2 M x = 0,
\]

\[
C^T x = 0,
\]

where \( \lambda (= k_0) \) is the eigenvalue and \( x \) is the eigenvector. The matrices \( A, R, M, \) and \( C \) in (3.15) have entries

\[
a_{ij} = \int_\Omega \mu^{-1}_r (\nabla \times N_i) \cdot (\nabla \times N_j) \, dx, \quad 1 \leq i, j \leq n,
\]

\[
m_{ij} = \int_\Omega \varepsilon_r N_i \cdot N_j \, dx, \quad 1 \leq i, j \leq n,
\]

\[
r_{ij} = i \int_\Gamma \sqrt{\varepsilon_r / \mu_r} (n \times N_i) \cdot (n \times N_j) \, ds + i \int_\Omega \sigma Z_0 N_i \cdot N_j \, dx, \quad 1 \leq i, j \leq n,
\]

\[
c_{i\ell} = \int_\Omega \varepsilon_r N_i(x) \cdot \nabla N_{\ell}(x) \, dx, \quad 1 \leq i \leq n,
\]

\[
1 \leq \ell \leq m.
\]

respectively. Since \( \varepsilon_r \) and \( \mu_r \) are element-wise constant, \( A, M, \) and \( R \) are composed of real symmetric element matrices multiplied by a complex factor. Thus, they are complex symmetric.

3.3.2 The eigensolver

Instead of solving the eigenvalue problem (3.15) we use an appropriate linearization \[74\]. Defining \( x_2 = \lambda x_1 \), (3.15) is transformed into a constrained linear eigenproblem of the form

\[
Ax \equiv \begin{bmatrix} A & O \\ O & I \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} R & M \\ I & O \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \equiv \lambda B x, \quad C^T x \equiv \begin{bmatrix} C \\ O \\ C \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.
\]

The identity matrix \( I \) in (3.17) could be replaced by any nonsingular matrix \( X \), i.e., we could write \( X x_2 = \lambda X x_1 \). Choosing \( X = M \) would make both \( A \) and \( M \) complex symmetric. Since complex symmetry does not necessarily increase numerical simplicity \[5\] we set \( X = I \).

The linearization of the quadratic eigenvalue problem allows us to use well-
known eigensolvers. For solving large sparse complex eigenvalue problems the Jacobi–Davidson QZ algorithm (JDQZ) is appropriate [84, 19]. JDQZ computes a partial QZ decomposition

\[ AQ = ZT_A, \quad BQ = ZT_B, \quad C^T Q = O, \]  

where \( Q \) and \( Z \) are \( 2n \times r \) matrices, and \( T_A \) and \( T_B \) are upper-triangular \( r \times r \) matrices. The quotients of corresponding diagonal elements of \( T_A \) and \( T_B \) provide the eigenvalues. We are looking for a few eigenvalues closest to a prescribed target value \( \tau \). The initial target is chosen to be an estimate of the expected eigenvalue. More details on selecting the initial target are given in Section 4.1. The columns of \( Q \) are the Schur vectors from which the desired eigenvectors are extracted. \( Z \) is a so-called shadow space. For details on eigenvalue/eigenvector extraction and on restarting we refer to [84]. Here, we only discuss the most time consuming step of the Jacobi–Davidson algorithm, the expansion of the search space. To that end, the so-called correction equation has to be solved,

\[ (I - \tilde{Z}^* \tilde{\lambda})(A - \tilde{\lambda}B)(I - \tilde{Q}\tilde{Q}^*) t = r \quad \tilde{Q}^* t = 0, \quad C^T t = 0, \]  

where the solution \( t \) is used to span the search space. \( \tilde{\lambda} \) is the actual best eigenvalue approximation, \( \tilde{Q} \) is the matrix \( Q \) expanded by the corresponding Schur vector \( u \). \( \tilde{Z} \) is the matrix \( Z \) expanded by the new shadow vector \( p \). So, \( \tilde{Q} \) and \( \tilde{Z} \) are \( 2n \times (r + 1) \) matrices.

The solution \( t \) of (3.19) is not needed to high accuracy. Therefore, we can approximately solve the correction equation by executing a few steps of a preconditioned Krylov space method. The Jacobi–Davidson method has the advantages of short execution time and low memory consumption, which is crucial for solving large scale eigenproblems. To impose the divergence-free condition in (3.15), we construct an appropriate projector to assert that each vector in the search space is in the null space of \( C^T \) [23, 3, 4].

The global matrix of the correction Eq. (3.19) has a \( 2 \times 2 \) block structure, cf. (3.17). We use the block Gauss–Seidel preconditioner on the global matrix. To solve with the diagonal (1,1) block, we use the Jacobi preconditioner or a hierarchical basis preconditioner [3, 4]. The diagonal (2,2) block is related to the identity matrix, so it is trivial to solve.

### 3.4 Validation and application of the algorithm

To validate the eigenmodal solver algorithm and, in particular, its combination with an absorbing boundary condition (ABC), we analyze three problems,
3 Plasmonic nanostructures

namely (i) the dielectric resonator antenna (DRA), (ii) the nano-cuboid, and (iii) a spherical nanoparticle with variable radius. Eventually, in addition to the benchmark cases, we apply the algorithm to the analysis of the optical dipole antenna.

Unless otherwise specified, quadratic Nédélec elements are used with each element having 20 local degrees of freedom. We use spatially highly resolved tetrahedral meshes. In particular, for the nano-optical application examples, the size of an element in the region of the resonating structure is in the order of a few nanometers, or smaller. In combination with second order basis functions we thus employ a highly accurate field approximation. Therefore, the error caused by the spatial discretization is generally negligible.

All simulations were carried out on the Cray XT5 at the Swiss National Supercomputing Centre (CSCS) [14]. We stress that femmaxx generally operates on any reasonable distributed memory parallel computer. Naturally, given the size of the numerical problem at hand, femmaxx indeed profits from increased compute power but it does by no means enforce the use of a supercomputer.

3.4.1 Dielectric resonator antenna

We analyze rectangular dielectric resonator antennas (DRAs) in the microwave region. DRAs have been analyzed theoretically [57, 50]. We study DRAs with various dimensions $a$, $b$, $d$, and dielectric function $\varepsilon_r$. The meshes usually contain $\approx 80'000$ tetrahedra for each DRA, so that the discretization counts $\approx 600'000$ degrees of freedom (dof).

The radius of the computational domain $\Omega$ is $R_b = 30$ mm, see Fig. 3.2(a). Theoretical solutions from [50] and numerical results for the $TE_{111}^z$ mode are given in Table 3.1. For all simulations, we set $\tau$ to a real value, the equivalent of 4.8 GHz. Each simulation completes in $\approx 3$ minutes using 64 cores on the Cray XT5. The normalized residual $(||Ax - \lambda Bx||_2/||x||_2)$ of the converged eigenpair is in the order of $10^{-5}$.

We note that the theory in [50] is based on a simplified model restricted to $\Omega_1$, i.e. the DRA, assuming perfect magnetic boundary conditions (PMC) $E \cdot n = 0$ on the surface $\Gamma_{13}$ of the DRA. In femmaxx $\Gamma_{13}$ is treated as an interface where just the continuity of the tangential component of $E$ is enforced. $E \cdot n$ need not be zero.

As expected, $Q_1$ matches very well with $Q_2$. However, the different treatment of the DRA walls causes a noticeable difference between $f$ and $f_{MI}$ and also between $Q_1$ and $Q_{MI}$. We note that, interestingly, our frequencies $f$ are closer to the experimental data cited in [50] than $f_{MI}$. Thus, the model implemented in femmaxx appears to be more realistic.

We select the last DRA as an example for the convergence behavior of
Table 3.1: $TE^{1}_{111}$ mode of DRAs ($f$: resonant frequency computed by Femaxx; $f_{MI}$: theoretically determined resonant frequency [50]; Diff.: relative difference of $f$ and $f_{MI}$; $Q_{MI}$: theoretically determined the Q factor [50].)

<table>
<thead>
<tr>
<th>$\varepsilon_r$</th>
<th>a (mm)</th>
<th>b (mm)</th>
<th>d (mm)</th>
<th>$f$ (GHz)</th>
<th>$f_{MI}$ (GHz)</th>
<th>Diff. (%)</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
<th>$Q_{MI}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>79.46</td>
<td>7.45</td>
<td>7.45</td>
<td>2.98</td>
<td>4.644</td>
<td>4.346</td>
<td>6.9</td>
<td>176.8</td>
<td>174.4</td>
<td>128.7</td>
</tr>
<tr>
<td>37.84</td>
<td>8.60</td>
<td>8.60</td>
<td>2.58</td>
<td>6.221</td>
<td>5.934</td>
<td>4.8</td>
<td>31.1</td>
<td>31.3</td>
<td>37.9</td>
</tr>
<tr>
<td>37.84</td>
<td>7.45</td>
<td>7.45</td>
<td>3.51</td>
<td>5.614</td>
<td>5.337</td>
<td>5.2</td>
<td>50.8</td>
<td>51.0</td>
<td>44.0</td>
</tr>
<tr>
<td>20.0</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>6.545</td>
<td>6.409</td>
<td>2.1</td>
<td>13.3</td>
<td>13.4</td>
<td>18.1</td>
</tr>
</tbody>
</table>

the solution w.r.t. the number $N$ of tetrahedral mesh elements. We let $N$ be 8’137, 28’150, 87’380 and 645’868, respectively. The computed resonant frequencies are, respectively, 6.548 GHz, 6.545 GHz, 6.545 GHz and 6.545 GHz, while the corresponding quality factors $Q_1$ are, respectively, 13.9, 13.4, 13.3 and 13.3. Thus, a tetrahedral mesh containing $\approx$ 80’000 elements for each DRA in Table 3.1 is good enough.

The electric field distribution $|\mathbf{E}|$ of the last DRA in Table 3.1 ($a = b = 10\text{ mm}, d = 4\text{ mm}$) is shown in Fig. 3.2(b). Here, the number of tetrahedra of the mesh is 645’868. According to the theoretical model in [50], the electric field inside the DRA along the $y$-axis satisfies $E_y = E_z = 0$ and $E_x = A k_y \sin(k_y y)$ where $A$ is a complex constant and $k_y = \pi/b$ is the wavenumber. We plot the three components of the electric field along the $y$-axis in Fig. 3.2(d). They are in good agreement with the theoretical model, except that $E_x$ varies a bit more than half a cycle inside the DRA. This, again, can be ascribed to the differing treatments of the DRA surface.

For DRA’s, in the microwave region of the electromagnetic spectrum, the $\varepsilon_r$ of the employed materials is generally constant. However, in the optical region of the electromagnetic spectrum the dielectric permittivity is in general not constant but depends on the real part of eigenvalue Re($k_0$), cf. (3.2). The dispersive material property implies that the matrices $M$ and $C$ in (3.16) also depend on Re($k_0$). So, in the nano-optical region, the eigenproblem (3.15) becomes a ‘truly’ non-linear, i.e. not linearizable, eigenproblem. At present, Femaxx is capable of solving linear and quadratic eigenproblems. Nevertheless, we can address this non-linear eigenproblem via solving a sequence of quadratic eigenproblems. We start with an estimate $\lambda(0)$ for the eigenvalue of the expected eigenmode. By means of Re($\lambda(0)$), we determine the dielectric permittivity $\varepsilon_r = \varepsilon_r(\text{Re}(\lambda(0)))$, based on [40], so that the matrices $M(\text{Re}(\lambda(0)))$, and $C(\text{Re}(\lambda(0)))$ can be constructed. Then, we solve the quadratic eigenvalue
3 Plasmonic nanostructures

(a) a 3-D sketch of the DRA (or nano-cuboid). (b): a DRA with $a = b = 10$ mm, $d = 4$ mm and $\varepsilon_r = 20.0$, resonance at 6.545 GHz. The electric field distribution $|\mathbf{E}|$ of the $TE_{111}$ mode in the antenna and its surrounding region is shown. (c): a gold cuboid with $a = b = 100$ nm, $d = 40$ nm, resonance at 935.1 THz (where $\varepsilon_r = -1.2308 + i5.8458$). It shows the electric field distribution $|\mathbf{E}|$ of the $TE_{111}$ mode in the cuboid and the surrounding region. (b) and (c) are visualized on the $xy$-plane through the center of the antenna. (d) 3 components of the electric field of the DRA (in (b)) plotted along the $y$-axis. (e) 3 components of the electric field of the gold cuboid (in (c)) plotted along the $y$-axis. Note that, $E_x$, $E_y$, and $E_z$ are complex. For $E_y$ and $E_z$ (approximately zero), only real parts are shown. $E_x$ is scaled such that $\|\text{Re}(E_x)\|_\infty = 1$. 

Figure 3.2: (a) a 3-D sketch of the DRA (or nano-cuboid). (b): a DRA with $a = b = 10$ mm, $d = 4$ mm and $\varepsilon_r = 20.0$, resonance at 6.545 GHz. The electric field distribution $|\mathbf{E}|$ of the $TE_{111}$ mode in the antenna and its surrounding region is shown. (c): a gold cuboid with $a = b = 100$ nm, $d = 40$ nm, resonance at 935.1 THz (where $\varepsilon_r = -1.2308 + i5.8458$). It shows the electric field distribution $|\mathbf{E}|$ of the $TE_{111}$ mode in the cuboid and the surrounding region. (b) and (c) are visualized on the $xy$-plane through the center of the antenna. (d) 3 components of the electric field of the DRA (in (b)) plotted along the $y$-axis. (e) 3 components of the electric field of the gold cuboid (in (c)) plotted along the $y$-axis. Note that, $E_x$, $E_y$, and $E_z$ are complex. For $E_y$ and $E_z$ (approximately zero), only real parts are shown. $E_x$ is scaled such that $\|\text{Re}(E_x)\|_\infty = 1$. 

40
3.4 Validation and application of the algorithm

problem (3.15) and consequently obtain an improved eigenvalue estimate $\lambda^{(1)}$. This procedure is repeated to yield further estimates $\lambda^{(2)}$, $\lambda^{(3)}$, etc. We terminate the iteration as soon as two consecutive estimates are close enough. In what follows, we use this iterative procedure. We first validate our method with the nano-cuboid and nano-sphere geometries. Then we use femaxx for the analysis of specific optical devices. In the individual quadratic eigenvalue problems we use $\tau = \text{Re}(\lambda^{(j)})$ as the target.

3.4.2 Cuboid

We investigate the $TE_{111}^z$ mode of a nano-cuboid and reuse the sketch in Fig. 3.2(a). We let $a = b = 100 \text{ nm}$, $d = 40 \text{ nm}$, $R_b = 300 \text{ nm}$ and assume gold and silver, respectively, for the material of the nano-cuboid. We also reuse the theoretical DRA model [57, 50] to compute the resonance frequency $f_{MI}$ and the electromagnetic field. The theoretically determined wavenumber $k_0$ is complex. The quality factor $Q_{MI}$ is also theoretically determined via Eq. (3.5). We comment that Mongia’s model was developed in the micro-region of the electromagnetic spectrum. Thus, strictly speaking, it may not be used as a reference or benchmark in the optical region where the dielectric permittivity of metals is highly dispersive and exhibits considerable dielectric loss. Nevertheless, we believe that, at present, there is no better analytical model available than Mongia’s for cuboid geometries. Therefore, we use Mongia’s model as a signpost into the optical region. Fully aware that we stretch Mongia’s model beyond its original limits of applicability, we consider it useful to help acquire at least qualitative insight into the mode structure of the nano-cuboid. The results are shown in Table 3.2. The mesh contains 645,868 tetrahedra leading to 4,120,676 degrees of freedom. For all simulations, the initial value of $\tau(= \text{Re}(\lambda^{(0)}))$ is equivalent of 1000 THz. The eigensolver calculation time for a single run is less than 10 minutes with 256 cores, while total computation time for one eigenmode is then below 40 minutes for 3–4 runs. The normalized residual of the converged eigenpair is around $10^{-5}$.

The electric field distribution $|E|$ of the gold cuboid is shown in Fig. 3.2(c). The three components of $E$, plotted along the $y$-axis, are displayed in Fig. 3.2(e). Again they are in good agreement with the fields of the theoretical model. In Fig. 3.2(e), $E_y \approx 0$ and $E_z \approx 0$. $E_x$ varies by a bit more than half a cycle inside the cuboid which indicates that the PMC is not strictly satisfied on the surfaces of the cuboid. We also notice that the field is strong inside the cuboid and drops to almost zero in the far-field zone, see Fig. 3.2(e). Therefore, the mode experiences rather high dissipative loss, but radiates only scant electromagnetic power. Thus, $\eta$ is very small.

We comment that the dielectric function $\varepsilon_r$ is not at all constant for gold and
3 Plasmonic nanostructures

Table 3.2: Numerical figures of merit for the $TE_{111}^z$ mode of the nano-cuboid where $\lambda$: resonant wavelength computed by femaxx; $f$: resonant frequency computed by femaxx; $f_{MI}$: theoretically determined resonant frequency [50]; Diff.: relative difference of $f$ and $f_{MI}$; $Q_{MI}$: theoretically determined Q factor [50].

<table>
<thead>
<tr>
<th>material</th>
<th>$\lambda$ (nm)</th>
<th>$f$ (THz)</th>
<th>$f_{MI}$ (THz)</th>
<th>Diff. (%)</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
<th>$Q_{MI}$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gold</td>
<td>320.6</td>
<td>935.1</td>
<td>739.8</td>
<td>26.4</td>
<td>0.47</td>
<td>0.82</td>
<td>0.39</td>
<td>5.9 $\times$ 10^{-6}</td>
</tr>
<tr>
<td>silver</td>
<td>200.7</td>
<td>1494</td>
<td>1107</td>
<td>35.0</td>
<td>0.56</td>
<td>0.79</td>
<td>0.55</td>
<td>2.5 $\times$ 10^{-7}</td>
</tr>
</tbody>
</table>

silver in the optical region of the spectrum. The dispersive material property, in combination with the different treatment of the DRA wall, largely explains the noticeable difference between $f$ and $f_{MI}$. Compared to dielectric resonator antennas, the quality factors ($Q_1$, $Q_2$, or $Q_{MI}$) of the nano-cuboid are very small, significantly below 1. This is attributed to high dissipative loss. The difference between $Q_1$ and $Q_2$ is noticeable, cf. the discussion in Section 3.2.

3.4.3 Sphere

We compute the resonant, plasmonic mode of a silver sphere, hovering in vacuum, with varying radius, i.e. $R = 25, 30, 35, 40, 60$ and $70$ nm. Thus, $\varepsilon_{sub} = \varepsilon_{env} = 1.0$. The radius of the computational domain is $R_b = 300$ nm. For each sphere, the mesh contains about 250'000 tetrahedra. For all simulations, the initial value of $\tau$ is equivalent of $3.00$ eV. In this example we had problems with the quadratic elements. The computation of $E_{res1}$ for $R = 25$ nm and of $E_{res2}$ (see below) for $R = 60$ nm failed to converge such that we had to resort to linear elements. In general, quadratic elements deliver more accurate solutions for fewer elements but they are more demanding w.r.t. iterative solver schemes. The eigensolver calculation time for a single run is less than 10 minutes with 256 cores (or 32 cores) if quadratic (or linear) elements are used. The normalized residual of the converged eigenpair is around $5 \times 10^{-4}$. The total computation time for one eigenmode is thus below 1 hour for 5–6 runs. We study the energy range from 2 eV to 4 eV, in wavelength units, 310 nm to 620 nm. We use $\lambda$ for the numerically determined resonance wavelength and the resonance energy $E = hc/\lambda$, with $h$ the Plank constant and $c$ the speed of light in vacuum, respectively.

For comparison, we also compute the theoretical Mie solution [10], for each radius and for the same energy range. When $R = 25, 30, 35$, or $40$ nm, the peaks of scattering ($Q_{sca}$) and absorption efficiency ($Q_{abs}$) occur at almost...
3.4 Validation and application of the algorithm

Figure 3.3: Mie solutions of silver spheres with radii of (a) 30 nm and (b) 60 nm.

the same position, $E_{\text{res1}}^{\text{mie}}$. In Fig. 3.3(a) we show the Mie solution for $R = 30$ nm. The full-linewidth-at-half-maximum (FWHM) energy width $\Delta E_{\text{ext}}$ of the extinction efficiency, $Q_{\text{ext}} = Q_{\text{sca}} + Q_{\text{abs}}$, can be evaluated. The quality factor is computed as $Q_{\text{res1}}^{\text{mie}} = E_{\text{res1}}^{\text{mie}} / \Delta E_{\text{ext}}$ [72].

When $R = 60$ or 70 nm, the peaks of $Q_{\text{sca}}$ and $Q_{\text{abs}}$ are separated from each other, such that there are two distinct resonances close together. Figure 3.3(b) shows the Mie solution for $R = 60$ nm. We first study the resonance $E_{\text{res1}}^{\text{mie}}$ where $Q_{\text{sca}}$ reaches maximum. Let $\Delta E_{\text{sca}}$ be the FWHM of $Q_{\text{sca}}$, and the quality factor of this resonance $Q_{\text{res1}}^{\text{mie}} = E_{\text{res1}}^{\text{mie}} / \Delta E_{\text{sca}}$. However, in such cases where two peaks overlap, $Q_{\text{sca}}$ deviates from a Lorentzian shape [20], so that evaluating $Q_{\text{res1}}^{\text{mie}}$ by FWHM $\Delta E_{\text{sca}}$ is not accurate.

The radiative quantum yield $\eta_{\text{mie}}$, computed with the Mie solution, is defined as the ratio between scattering and extinction efficiency, i.e., $\eta_{\text{mie}} = Q_{\text{sca}} / Q_{\text{ext}}$.

In Table 3.3 and Fig. 3.4, we show the numerical results, compared with Mie solutions. In general, the resonance ($E_{\text{res1}}$ and $E_{\text{res1}}^{\text{mie}}$) and radiative quantum yield ($\eta$ and $\eta_{\text{mie}}$) show good agreements. The differences are mainly attributed to the 1st order ABC employed by femaxx, while the Mie model is unbounded. Quality factors ($Q_1$ and $Q_{\text{res1}}^{\text{mie}}$) also match when $R = 25, 30, 35$ or 40 nm. When $R = 60$ or 70 nm, two peaks of resonances ($E_{\text{res1}}^{\text{mie}}$ and $E_{\text{res2}}^{\text{mie}}$) exist and are close together, $Q_{\text{res1}}^{\text{mie}}$ is not accurate, see discussion above, and $Q_1$ is more reliable. As noted in Section 3.2, in the range of 2.5 eV to 3.5 eV, the difference of silver $\varepsilon_r$ between the Drude model and the experimental data is noticeable. Therefore $Q_2$ is not accurate either. Note that $E_{\text{res1}}$ approximates
Table 3.3: Results of the numerical analysis and the semi-analytical Mie solutions of the silver sphere hovering in vacuum, where: $E_{\text{res}1}$: resonant energy computed by \texttt{femaxx}; $E_{\text{diff}}$: the relative difference between $E_{\text{res}1}$ and $E_{\text{res}1}^{\text{mie}}$, $Q_{\text{diff}}$: the relative difference between $Q_1$ and $Q_{\text{res}1}^{\text{mie}}$.

<table>
<thead>
<tr>
<th>$R$ (nm)</th>
<th>$\lambda$ (nm)</th>
<th>$E_{\text{res}1}$ (eV)</th>
<th>$E_{\text{res}1}^{\text{mie}}$ (eV)</th>
<th>$E_{\text{diff}}$ (%)</th>
<th>$\eta$</th>
<th>$\eta_{\text{mie}}$</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
<th>$Q_{\text{res}1}^{\text{mie}}$</th>
<th>$Q_{\text{diff}}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>368.7</td>
<td>3.36</td>
<td>3.41</td>
<td>1.5</td>
<td>0.57</td>
<td>0.48</td>
<td>21.5</td>
<td>23.2</td>
<td>22.7</td>
<td>5.3</td>
</tr>
<tr>
<td>30</td>
<td>373.4</td>
<td>3.32</td>
<td>3.37</td>
<td>1.5</td>
<td>0.64</td>
<td>0.63</td>
<td>17.0</td>
<td>20.7</td>
<td>16.9</td>
<td>0.6</td>
</tr>
<tr>
<td>35</td>
<td>381.0</td>
<td>3.25</td>
<td>3.33</td>
<td>2.4</td>
<td>0.78</td>
<td>0.75</td>
<td>10.1</td>
<td>19.3</td>
<td>12.3</td>
<td>17.9</td>
</tr>
<tr>
<td>40</td>
<td>388.4</td>
<td>3.19</td>
<td>3.25</td>
<td>1.9</td>
<td>0.81</td>
<td>0.86</td>
<td>9.6</td>
<td>17.1</td>
<td>8.8</td>
<td>9.1</td>
</tr>
<tr>
<td>60</td>
<td>408.9</td>
<td>3.03</td>
<td>3.02</td>
<td>0.3</td>
<td>0.88</td>
<td>0.95</td>
<td>6.0</td>
<td>12.3</td>
<td>3.0</td>
<td>100</td>
</tr>
<tr>
<td>70</td>
<td>440.6</td>
<td>2.81</td>
<td>2.81</td>
<td>0</td>
<td>0.95</td>
<td>0.98</td>
<td>5.3</td>
<td>8.6</td>
<td>2.1</td>
<td>152</td>
</tr>
</tbody>
</table>

a 3-fold degenerate resonance. The corresponding computed mode splits in three simple modes since the mesh does not respect the spherical symmetry of the problem. For $R = 60$ nm, the relative differences among the three computed resonances are below 1%.

We use $R = 60$ nm as an example (see Fig. 3.3(b)), and study the resonance $E_{\text{res}2}$ where $Q_{\text{abs}}$ reaches the maximum. (Our computations indicate that $E_{\text{res}2}$ is 5-fold degenerate.) The numerical value of this mode is $E_{\text{res}2} = 3.48$ eV and $Q_1 = 8.1$. Let the angular frequency, dissipated energy and radiated powers of resonance $E_{\text{res}2}$ be $\omega_{\text{res}2}$, $U_d^{\text{res}2}$ and $P_r^{\text{res}2}$, respectively. Further, let the corresponding quantities for resonance $E_{\text{res}1}$ be $\omega_{\text{res}1}$, $U_d^{\text{res}1}$ and $P_r^{\text{res}1}$. The electric fields are scaled so that $\int_{\Omega} \text{Re}(\varepsilon_r)|\mathbf{E}(x)|^2 \, dx$ are equal. We define $r_d = (\omega_{\text{res}2}U_d^{\text{res}2})/(\omega_{\text{res}1}U_d^{\text{res}1})$ and $r_r = P_r^{\text{res}2}/P_r^{\text{res}1}$. Then $r_d = 8.9$ and $r_r = 0.21$. Resonance $E_{\text{res}2}$ has higher dissipative loss since $Q_{\text{abs}}$ reaches maximum, while resonance $E_{\text{res}1}$ has higher radiation loss since $Q_{\text{sca}}$ reaches maximum. The Mie solution of this mode is $E_{\text{res}2}^{\text{mie}} = 3.45$ eV. Again, evaluating $Q_{\text{res}2}^{\text{mie}} (= 17.2)$ by FWHM $\Delta E_{\text{abs}}$ is not accurate. The electric field plots of the two modes, when $R = 60$ nm, are shown in Figures 3.5(a) and 3.5(b).

We use the resonance $E_{\text{res}1}$ of the silver sphere ($R = 60$ nm) to study the influence of the 1st order ABC on the accuracy of the result. If we let the boundary radius $R_b$ vary from 250 to 300 nm (keeping the number of tetrahedra around 250000), the deviations of the computed resonant wavelength are as small as 3 nm. For $R_b = 300$ nm, the computed resonant wavelength (408.9 nm) is closest to the resonant wavelength of the Mie solution ($\approx 410.5$ nm). For smaller $R_b$, the error gets larger. For $R_b = 200$ nm, e.g., the computed resonant wavelength is 385.0 nm.
3.4 Validation and application of the algorithm

Figure 3.4: Numerical and Mie solutions of Q factors and radiative quantum yield (subfigure) of silver spheres with varying radius $R$.

![Graph showing numerical and Mie solutions of Q factors and radiative quantum yield with varying radius $R$.]

Figure 3.5: (a) and (b) show the electric field distributions ($|E|$) in the vicinity of a silver sphere on the $xy$-plane. The radius of the sphere is 60 nm. (a): the mode when $Q_{\text{sca}}$ reaches maximum, $E_{\text{res}} = 3.03$ eV; (b): the mode when $Q_{\text{abs}}$ reaches maximum, $E_{\text{res}} = 3.48$ eV.
3.4.4 An optical dipole antenna

We study the electromagnetic near-field of the gold nano-optical dipole antenna investigated in [41]. The geometry is shown in Fig. 3.6(a) and a 3-D sketch is in 3.6(b). The antenna has two arms, each of which has dimensions \( a = b = 40 \, \text{mm}, \, l = 100 \, \text{nm} \). The corners of each arm are rounded, with a radius of curvature \( r = 5 \, \text{nm} \). The gap \( g \) between the two arms is 20 nm wide but will be varied in order to study the electromagnetic field enhancement as a function of the gap width. The radius \( R_b \) of the spherical computational domain is 400 nm.

We study three different dielectric material arrangements: (1) the antenna hovers in vacuum, thus \( \varepsilon_{\text{sub}} = \varepsilon_{\text{env}} = 1.0 \); (2) the antenna resides on a silica substrate, thus \( \varepsilon_{\text{sub}} = 2.25, \varepsilon_{\text{env}} = 1.0 \); (3) the antenna resides on a silica substrate and is covered with water, thus \( \varepsilon_{\text{sub}} = 2.25, \varepsilon_{\text{env}} = 1.77 \) [32]. We understand this arrangement to be a relatively detailed model for an optical antenna used in biochemical sensing where many processes are studied in aqueous solutions.

We start by considering a single arm in vacuum whose resonant wavelength is 623.5 nm. When bringing two such arms into close proximity, the interaction of the two fundamental eigenmodes lead to mode splitting. When these two arms come close enough to each other, the so-called bright, aka. optically active, mode and the dark mode split. The modes can then be well distinguished in the scattering spectra [12, 83].

Figure 3.6: The nano-optical dipole antenna. (a) geometry; (b) 3-D sketch.
3.4 Validation and application of the algorithm

**Bright mode**

We compute the bright mode within the spectral range of 400 nm to 1000 nm. For all simulations, the initial value of $\tau$ is equivalent of 660 nm. The results are presented in Table 3.4.

![Image](image_url)

(a)

Figure 3.7: (Numerical analysis of a gold, nano-optical dipole antenna (model (1)). The gap width is 20 nm and the resonance is at 665.7 nm; (a) magnetic and (b) electric field distribution ($|\mathbf{H}|$ and $|\mathbf{E}|$) in the vicinity of the antenna surface; (c) electric field distribution $|\mathbf{E}|$ around the gap; (a), (b), and (c) are evaluated on the $xy$-plane through the center of the antenna arms.

The fields of model (1), an antenna in vacuum, are evaluated in the sample domain $\Omega_{\text{sample}} = 400 \text{ nm} \times 200 \text{ nm} \times 200 \text{ nm}$. Both magnetic and electric fields are shown in Fig. 3.7(a) and 3.7(b), respectively. A close-up view of the electric field in the gap region is shown in Fig. 3.7(c). The results in Table 3.4 confirm the influence of the substrate onto the antenna’s resonant wavelength. With increasing permittivity $\varepsilon_{\text{sub}}$ the antenna resonance is red-shifted; with increasing $\varepsilon_{\text{env}}$ the resonance wavelength is also red-shifted [18]. Each single
3 Plasmonic nanostructures

Table 3.4: Numerical analysis of a gold nano-optical antenna (#p: number of cores; \( N \): number of tetrahedra; dofs: degrees of freedom; \( \lambda \): resonant wavelength computed by \texttt{femmax}; \( Q_1 \) and \( Q_2 \): quality factors by Eq. (3.5) and Eq. (3.10), respectively;

<table>
<thead>
<tr>
<th>model</th>
<th>( \varepsilon_{\text{sub}} )</th>
<th>( \varepsilon_{\text{env}} )</th>
<th>#p</th>
<th>( N )</th>
<th>dofs</th>
<th>( \lambda ) (nm)</th>
<th>( Q_1 )</th>
<th>( Q_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) in vacuum</td>
<td>1.0</td>
<td>1.0</td>
<td>300</td>
<td>340'269</td>
<td>2'162'346</td>
<td>665.7</td>
<td>20.0</td>
<td>19.8</td>
</tr>
<tr>
<td>(2) on substrate</td>
<td>2.25</td>
<td>1.0</td>
<td>600</td>
<td>597'408</td>
<td>3'791'532</td>
<td>717.5</td>
<td>12.4</td>
<td>13.6</td>
</tr>
<tr>
<td>(3) on substrate, water-covered</td>
<td>2.25</td>
<td>1.77</td>
<td>600</td>
<td>597'408</td>
<td>3'791'532</td>
<td>801.7</td>
<td>10.6</td>
<td>14.3</td>
</tr>
</tbody>
</table>

Figure 3.8: plots the electric field amplitudes \( |E| \) along the \( x \)-axis, associated with 3 material arrangements.

run of the quadratic eigensolver requires 10 to 15 minutes. The total computation time for one bright mode is then below 90 minutes for 5–6 runs. The normalized residual of the converged eigenpair is around \( 5 \times 10^{-4} \). The memory per core required by the eigensolver is 33.38 MB for the model (1), and 28.65 MB for models (2) and (3). Thus, the method is particularly attractive when compared to other 3-dimensional FEM approaches [70, 43]. From Figs. 3.7(b) and 3.7(c) we note that a high-intensity electric field is generated in the vicinity of the antenna surface. It is stronger in the gap (Fig. 3.8), and it is strongest at the inner corners of the antenna.

We now investigate how the gap width \( g \) affects the resonance wavelength \( \lambda \) and the field intensity \( |E|^2 \) in the gap. We use the dielectric arrangement of model (2). Let the gap width \( g \) be 20 nm, 10 nm, 5 nm, and 1 nm, respectively. The corresponding resonance wavelengths are then 717.5 nm, 751.9 nm, 796.3 nm, and 913.5 nm, respectively. A red shift, i.e., increasing wavelength,
3.4 Validation and application of the algorithm

Figure 3.9: Numerical analysis of a gold nano-optical antenna (model (2)): (a) gap width = 5 nm, resonance at 796.3 nm. It shows the electric field distribution $|E|$ (visualized on the $xy$-plane) in the vicinity of the antenna surface; (b) plots of the electric field amplitude $|E|$ along the $x$-axis with varying gap width $g$.

of the antenna resonance occurs with decreasing gap widths. Since the gap widths are significantly smaller than the resonance wavelengths, we can approximate the gap region with an electrostatic regime. Thus, we understand the gap region, including those portions of the dipole antenna, that immediately border it, to form a plate capacitor whose capacitance is given as $C = \varepsilon_0 \varepsilon_r A/d$. Here, $A$, $d$, and $\varepsilon_r$ correspond to the plate area, the distance between the plates and the dielectric property of the medium between the plates, respectively; in particular, $\varepsilon_r = 1$ since the gap is in vacuum. Then, it becomes clear that, with decreasing gap width, capacitance increases. A capacitance increase on the other hand is roughly proportional to a higher value of the dielectric permittivity and thus, for constant length of the dipole
arms, the resonance wavelength increases since the resonance frequency is reduced. We comment that adding a capacitor parallel to an antenna’s terminal contacts is a well-known technique in microwave electronics in order to geometrically shorten the antenna, thus mimicking higher dielectric permittivity in the vicinity of the antenna.

On the other hand, when decreasing the gap width $g = 10 \text{ nm}, 5 \text{ nm}, \text{ and } 1 \text{ nm}$, the corresponding field intensities $|E|^2$ increase and are $3.4, 9.6, \text{ and } 56$ times larger than when a gap width $g = 20 \text{ nm}$ is employed, see Fig. 3.9(b). Decreasing the gap width thus dramatically increases the field intensity in the gap. An example field distribution, for a $5 \text{ nm}$ gap width, is shown in Fig. 3.9(a). Due to the substrate effect the field is not symmetric with respect to the $x$-axis.

**Dark mode**

We study the dark mode of the dipole antenna in vacuum, model (1). The respective charge profiles [12] of the bright and the dark modes are shown in Fig. 3.10(a). Let the angular frequency, dissipative energy and radiated powers of the dark mode be $\omega_{\text{dark}}, U_d^{\text{dark}} \text{ and } P_r^{\text{dark}}$, respectively; further, let the corresponding quantities for the bright mode be $\omega_{\text{bright}}, U_d^{\text{bright}} \text{ and } P_r^{\text{bright}}$. Again, the electric fields are scaled so that the integrals of the type $\int_{\Omega} \text{Re}(\varepsilon_r)|E(x)|^2 \text{d}x$ are equal. We define $r_d = (\omega_{\text{dark}} U_d^{\text{dark}})/(\omega_{\text{bright}} U_d^{\text{bright}})$ and $r_r = P_r^{\text{dark}}/P_r^{\text{bright}}$. The radiative quantum yield $\eta$ is computed via Eq. (3.11). Associated results are listed in Table 3.5. The mesh contains about 350'000 tetrahedra which results in more than 2’000’000. For all simulations of the dark modes, the initial value of $\tau$ is equivalent of $550 \text{ nm}$. The eigensolver calculation time for a single run is 10 to 15 minutes with 256 cores, and the normalized residual of the converged eigenpair is around $5 \times 10^{-4}$ for the bright mode and $10^{-4}$ for the dark mode. The total computation time for one eigenmode is then below 90 minutes for 5–6 runs.

<table>
<thead>
<tr>
<th>$g$ (nm)</th>
<th>$\lambda$(nm)</th>
<th>$\eta$</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
<th>$\lambda$(nm)</th>
<th>$\eta$</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
<th>$r_d$</th>
<th>$r_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>582.0</td>
<td>$2.4 \times 10^{-3}$</td>
<td>0.89</td>
<td>13.3</td>
<td>665.7</td>
<td>0.27</td>
<td>20.0</td>
<td>19.8</td>
<td>10.1</td>
<td>0.067</td>
</tr>
<tr>
<td>10</td>
<td>557.4</td>
<td>$7.0 \times 10^{-4}$</td>
<td>0.78</td>
<td>9.7</td>
<td>689.8</td>
<td>0.29</td>
<td>23.3</td>
<td>20.3</td>
<td>18.2</td>
<td>0.032</td>
</tr>
<tr>
<td>5</td>
<td>532.1</td>
<td>$2.3 \times 10^{-4}$</td>
<td>0.65</td>
<td>7.4</td>
<td>713.9</td>
<td>0.26</td>
<td>21.5</td>
<td>20.7</td>
<td>44.9</td>
<td>0.029</td>
</tr>
</tbody>
</table>

When the width $g$ of the gap between the two arms is reduced, the single arm’s degenerate fundamental mode with its resonance at $623.5 \text{ nm}$ splits into...
3.5 Discussion and conclusions

A bright and a dark mode. With decreasing gap width, the bright mode redshifts, due to the effect of increased capacity for smaller width, and the dark mode shifts to blue. The dark mode radiates less, i.e. \( r_r < 0.1 \), which implies that it cannot efficiently radiate energy into the far-field, thus earning its dark reputation. On the other hand, the dark mode experiences higher dissipative loss inside the two gold arms, i.e. \( r_d > 10 \). Its dissipative loss approximates almost 100% of the total loss, namely \( \eta < 0.005 \), and thus the quality factor \( Q_1 \) of the dark mode is smaller than the bright mode.

The resonances of all dark modes are above 2 eV, i.e. below 620 nm; on the other hand, the resonances of all bright modes are all below 2 eV, i.e. above 620 nm. Therefore, computing the stored energy with Eq. (3.7) is a good approximation for bright modes, but significantly less so for dark modes, cf. the discussion in section 3.2. As a consequence, \( Q_2 \) is in good agreement with \( Q_1 \) for bright modes, but is significantly off for dark modes.

An exemplary field distribution for the dark mode, associated with a 10 nm gap width, is shown in Fig. 3.10(b). In Fig. 3.10(c) the field plots of bright and dark modes for a 5 nm gap width are displayed. The field associated with a dark mode is not only generated in the vicinity of the antenna surface, but also inside the gold arms, corresponding to a large \( U_{d \text{dark}} \). The field intensity is quite low in the gap.

3.5 Discussion and conclusions

We have implemented a 3-dimensional finite element method for the full-wave numerical analysis of electromagnetic resonators, surrounded by free-space. Specifically, there are no cavity walls, and the resonator structures need not be enclosed within perfect electric conductor (PEC) boundary conditions. Instead, we need to apply an absorbing boundary condition (ABC) scheme. The associated quadratic eigenvalue problem is linearized so that we can employ our current implementation of the Jacobi–Davidson QZ algorithm.

We have analyzed four different electromagnetic problems. The dielectric resonator antenna (DRA), analyzed in the microwave region of the spectrum, clearly demonstrates that the proposed Jacobi–Davidson algorithm is capable of calculating the dominant mode with excellent accuracy. The differences between the theoretically determined resonance frequencies, electric field and Q factor and their numerical counterparts are attributed to the fact that, in the analytical model, the perfect magnetic conductor (PMC) condition, is imposed on all surfaces of DRA, while femaxx does not need this simplifying assumption. The theoretical model of the DRA is reused as a signpost into the optical region for simulations of a nano-cuboid. The \( TE_{111}^z \) mode of the nano-cuboid
3 Plasmonic nanostructures

Figure 3.10: (a) Charge profiles; (b) the dark mode of the gold nano-optical antenna (gap width $g = 10$ nm) in vacuum: resonance at 557.4 nm. The electric field distribution $|\mathbf{E}|$ (visualized on the $xy$-plane) in the vicinity of the antenna surface is shown; (c) the gold nano-optical antenna (gap width $g = 5$ nm) in vacuum. The electric field amplitude $|\mathbf{E}|$ along the $x$-axis is plotted, associated with the bright mode (resonance at 713.9 nm) and the dark mode (resonance at 532.1 nm), respectively.
has a field pattern similar to the DRA, but exhibits a rather small quality factor due to high dissipative loss in metals. There is a noticeable difference between the numerically determined resonance wavelength and quality factor and the analytically determined resonance wavelength and quality factor, respectively. On the other hand, the electromagnetic field solutions demonstrate amazing agreement, in particular when we consider that Mongia’s model has been stretched considerably beyond its limits of applicability. The analysis of electromagnetic scattering off spheres represents a widely used benchmark in nano-optics which a computational electrodynamics code must be able to solve. In general, we observe good agreement between the theoretically determined resonance frequency, radiative quantum yield and Q factor and their numerical counterparts.

Thus, now that femaxx has passed a series of benchmarks, we apply it to more complex geometries and configurations. One of them, the optical dipole antenna, represents an important, elementary building block in many detection and sensing devices, especially for the life and materials sciences. We observe excellent agreement with previously published results by Kern and Martin [41], who simulate a dipole antenna with 20 nm gap width. We emphasize that our algorithm allows for the simple inclusion of background media and also media covering the antenna, e.g. an aqueous solution covering the antenna. This corresponds to a setup that is representative for many experimental arrangements. We note that the surface integral equation (SIE) method is also capable of modeling background media but at the cost of considerably increased complexity since it must include additional Green’s functions to accommodate, e.g., a substrate. Meanwhile, the finite element method ‘absorbs’ all this additional geometrical and material complexity into the corresponding mesh. On the other hand, the finite element method needs a transparent boundary condition of high-quality which should be placed as close as possible to the resonating structure, thus ensuring that precious mesh elements are not wasted but are put to good use when modeling fine geometrical details. We also note that the dark mode, with little radiated power, is another ideal test for our approach. A precise understanding of the dark modes is important for plasmonic applications [83]. Indeed, we have found the dark mode, starting from the original single arm configuration with a degenerate fundamental mode, which also strengthens the confidence into our solver. In conclusion, our approach is flexible enough to analyze nano-optical devices with almost arbitrary geometry or material arrangements. The femaxx code is therefore a useful instrument for the analysis of technologically relevant nano-optical device concepts.
Chapter 4

A nonlinear eigensolver for the analysis of dispersive nanostructures

Here, we introduce the electromagnetic eigenmodal solver code FemaxxNano for the numerical analysis of nanometer structured optical systems [27], a scientific field generally known as nanooptics. FemaxxNano solves the electric field vector wave equation and calculates the electromagnetic eigenmodes of nearly arbitrary 3-dimensional resonators, embedded either in free-space, vacuum or a background medium. Here, the study of the interaction between nanometer sized metallic structures and light is at the heart of the physical problem. Since metals in the optical region of the electromagnetic spectrum are highly dispersive and, thus, dissipative, dielectric media, we eventually obtain a non-linear eigenvalue problem. We discretize the electromagnetic eigenvalue problem with the finite element method (FEM) in 3-dimensional space and on unstructured tetrahedral grids. We introduce a fully iterative scheme to solve the non-linear problem for complex coefficient matrices that depend on wavelength. We investigate the properties of the algorithm in detail and demonstrate its performance by analyzing a nanometer sized optical dimer structure, a specific type of optical antenna, on distributed-memory parallel computers.

4.1 Introduction

The experimental and theoretical study of the interaction of nanometer sized metallic and dielectric structures with light is generally known as the field of nanooptics, see, e.g., [55] for an introduction. If the interaction is tied to metals, the field is often denoted plasmonics [1, 47, 37]. Plasmonic systems and devices are especially promising concepts in biosensing [1, 47, 37] and photovoltaics [6, 66]. In particular, theoretical plasmonics is concerned with the detailed calculation of the electromagnetic field distribution when a sub-wavelength sized, metallic particle interacts with light. Then, the be-
behavior of the electromagnetic system is rarely \emph{a priori} intelligible and often counter-intuitive. The analytical calculation of the field distribution is generally restricted to rather simple geometrical shapes illuminated by a transverse-electric-magnetic (TEM) wave, such as spheres, a problem known as Mie scattering, ellipsoids or hexahedral objects \cite{10}. In order to understand the mechanism of a specific plasmonic configuration, to study the sensitivity of a plasmonic biosensor or to compute the efficiency of a novel concept for a solar cell, the numerical analysis of the geometry at hand is therefore indispensable. In addition, since the fabrication of plasmonic devices is often expensive and time-consuming, a robust and reliable, pre-fabrication numerical simulation is invaluable. Therefore, a considerable body of literature on numerical methods for the computational analysis of plasmonic systems has been published, see \cite{70, 43, 41, 16} among many other excellent studies.

In a previous publication \cite{29}, we have introduced a novel technique for the numerical analysis of nanometer sized structures. The technique has certain advantages when compared to other methods. In particular, it is flexible enough to analyze nanostructures of almost arbitrary geometry and material arrangements. We have discretized the 3-dimensional electric field vector wave equation in the time-harmonic regime with the finite element method. We consider material losses through the complex permittivity, also known as dielectric function, \( \varepsilon_r \).

Since the finite element approach is used to discretize a volume, we need a boundary condition in order to truncate the computational domain such that power, radiated away from the resonating structure by an electromagnetic eigenmode, does not experience artificial reflection back into the computational domain. In other words, we need a transparent boundary condition \cite{79} that allows us to model a nanostructure, situated in free-space. This is rather different from the situation commonly encountered in the microwave frequency domain where resonating cavities employ a metallic boundary as an essential part of their design \cite{63}. In \cite{29}, and also in this study, we employ the first order absorbing boundary condition (ABC) \cite{63, 39}, that leads to a quadratic eigenvalue problem. We compute the electromagnetic modes associated with resonant nanostructure using a variant of the Jacobi–Davidson algorithm \cite{29}.

In this study we consider the dispersive, i.e. wavelength dependent, dielectric permittivity of metals in the optical region of the electromagnetic spectrum. We derive and solve the corresponding nonlinear electromagnetic eigenmodal problem,

\[
T(\lambda)x = Ax + \lambda Rx - \lambda^2 M(\text{Re}(\lambda))x = 0, \quad x \neq 0,
\]

(4.1a)
4.2 Formulation of the problem

with the constraint, i.e. the divergence-free condition

\[ C(\text{Re}(\lambda))^T \mathbf{x} = \mathbf{0}. \]  

(4.1b)

Here, \( \lambda \) is the eigenvalue, \( \text{Re}(\lambda) \) is its real part, and \( \mathbf{x} \) is the eigenvector. While matrices \( A \) and \( R \) are constant, \( M \) and \( C \) depend on \( \text{Re}(\lambda) \). Since solving (4.1) is a critical step in our technique to analyze optical nanostructures, a reliable and robust solver becomes vitally important. On the other hand, an efficient and reliable solver for Eqs. (4.1) remains a challenge. To the best of our knowledge we are not aware of a parallel solver for such large scale 3-dimensional problems. In this study, we introduce a parallel nonlinear eigensolver, \texttt{FemaxxNano}, for solving (4.1). The algorithm is parallelized for distributed memory machines, using the Trilinos framework [35].

The paper is organized as follows. In Section 4.2, we introduce the problem and its finite element discretization. In Sections 4.3, we investigate the algorithm that is implemented in \texttt{FemaxxNano}. In Section 4.4, we use examples from [69] to demonstrate capability and performance of \texttt{FemaxxNano}. We draw our conclusions in Section 4.5.

4.2 Formulation of the problem

Here, we briefly introduce the formulation for the electromagnetic eigenmodal analysis of optical nanostructures. For a detailed overview of the formulation we refer to [29]. We employ a bounded computational domain \( \Omega = \Omega_1 \cup \Omega_2 \), cf. Fig. 4.1, that encloses the nanostructure \( \Omega_1 \) and the background medium \( \Omega_2 \). We model the metallic loss of the nanostructure \( \Omega_1 \) with the complex permittivity \( \varepsilon_r \), often with a negative real part for metals in the optical region of the spectrum. The dielectric function \( \varepsilon_r \) depends on the angular frequency \( \omega \), or, equivalently, the photon energy \( = \hbar \omega \) where \( \hbar \) is the reduced Planck constant. For \( \varepsilon_r \) as a function of \( \omega \) we use the experimental data provided by Johnson and Christy [40] and employ piecewise linear interpolation. In Section 5, it is assumed that \( \Omega_2 \) is vacuum, therefore \( \mu_r = \varepsilon_r = 1 \) in \( \Omega_2 \). For the nanostructure \( \Omega_1 \), \( \mu_r = 1 \), where \( \varepsilon_r \) depends on \( \omega \), i.e.,

\[
\varepsilon_r = \begin{cases} 
\varepsilon_r(\omega), & \mathbf{x} \in \Omega_1 \\
1, & \mathbf{x} \in \Omega_2 
\end{cases}
\]  

(4.2)

We also define a parameter \( D = \omega \frac{d\varepsilon_r}{d\omega} \) to measure the dispersion. If \( D \) increases, so does the metallic dispersion, and vice versa; cf. Fig. 4.2 which plots \( \varepsilon_r \) and \( D \) for gold.
In the time-harmonic regime, the electric field $\mathbf{E}(x)$ obeys [29]

$$\nabla \times (\mu_r^{-1} \nabla \times \mathbf{E}(x)) - k_0^2 \varepsilon_r \mathbf{E}(x) = 0, \quad \nabla \cdot (\varepsilon_r \mathbf{E}(x)) = 0, \quad x \in \Omega. \tag{4.3}$$

Here, $k_0 = \frac{\bar{\omega}}{c}$ is the complex wavenumber in free space, $\bar{\omega} = \omega + i\alpha$ is the complex angular frequency, $\omega$ is the angular frequency, $\alpha$ is the exponential decay rate and $c$ is the speed of light in vacuum. We note that $\omega = \text{Re}(k_0)c$; $\mu_r$ and $\varepsilon_r$ are magnetic relative permeability and electric relative permittivity, respectively. On the truncation surface of the computational domain we employ the first order absorbing boundary condition (ABC) [63, 39]

$$n \times \nabla \times \mathbf{E}(x) = -ik n \times (n \times \mathbf{E}(x)). \tag{4.4}$$

Here, $k = k_0\sqrt{\mu_r\varepsilon_r}$ is the wavenumber of the surrounding medium. A natural weak form of (4.3)–(4.4) is then [29]

Find $k_0 \in \mathbb{C}$ and $\mathbf{E} \in V$, $\mathbf{E} \neq \mathbf{0}$, such that for all $\mathbf{f} \in V$ and all $q \in W$

$$\int_{\Omega} \mu_r^{-1} \nabla \times \mathbf{f} \cdot \nabla \times \mathbf{E} \, dx + ik_0 \int_{\Gamma} \frac{\varepsilon_r}{\mu_r} (n \times \mathbf{f}) \cdot (n \times \mathbf{E}) \, dx - k_0^2 \int_{\Omega} \varepsilon_r \mathbf{f} \cdot \mathbf{E} \, dx = 0, \tag{4.5}$$

$$\int_{\Omega} \varepsilon_r \mathbf{E} \cdot \nabla q \, dx = 0.$$
4.2 Formulation of the problem

Here, \( V \) denotes the functions in \( H(\text{curl}; \Omega) \) that satisfy the boundary condition (4.4) and \( W = H^1_0(\Omega) \). For details on these function spaces see [25, 39, 51]. We discretize (4.2) by the Ritz-Galerkin method [39] employing appropriate finite element subspaces of \( V \) and \( W \). To that end we decompose \( \Omega \) into tetrahedra. The triangulation respects the interface \( \Gamma_{12} \) such that \( \varepsilon_r \) is constant within each tetrahedron. The electric vector functions in \( V \) are approximated by Nédélec edge elements, while the scalar functions in \( W \) are approximated by Lagrange (nodal) finite elements [39]. This approach avoids the generation of spurious eigensolutions, and imposing the boundary conditions is straightforward [39]. Let the vector functions \( N_i, 1 \leq i \leq n \), be the Nédélec basis functions, while the scalar functions \( N_\ell, 1 \leq \ell \leq m \), denote the Lagrange basis functions. Then, the equations in (4.5) yield a constrained complex nonlinear eigenvalue problem.

\[
\begin{align*}
T(\lambda)x &= Ax + \lambda Rx - \lambda^2 M(\text{Re}(\lambda))x = 0, \\
C(\text{Re}(\lambda))^T x &= 0,
\end{align*}
\]  

(4.6)

where \( \lambda(= k_0) \) is the eigenvalue and \( x \) is the eigenvector. The matrices \( A, R, M, \) and \( C \) in (4.6) have entries

\[
\begin{align*}
a_{ij} &= \int_\Omega \mu_r^{-1}(\nabla \times N_i) \cdot (\nabla \times N_j) \, dx, & 1 \leq i, j \leq n, \\
m_{ij} &= \int_\Omega \varepsilon_r N_i \cdot N_j \, dx, & 1 \leq i, j \leq n, \\
r_{ij} &= i \int_\Gamma \sqrt{\varepsilon_r/\mu_r} (n \times N_i) \cdot (n \times N_j) \, dx, & 1 \leq i, j \leq n, \\
c_{i\ell} &= \int_\Omega \varepsilon_r N_i \cdot \nabla N_\ell \, dx, & 1 \leq i \leq n, \quad 1 \leq \ell \leq m.
\end{align*}
\]  

(4.7)

\( A \) is a real symmetric positive-semidefinite matrix. The entries \( r_{ij} \) are integrals on the boundary \( \Gamma \) where \( \varepsilon_r = 1 \), cf. (4.2). Therefore, \( R \) is a purely-imaginary symmetric matrix. The complex symmetric matrix \( M \) depends on \( \text{Re}(\lambda) \). The dielectric function \( \varepsilon_r \) in \( \Omega_1 \) differs from that in \( \Omega_2 \). Therefore, it is impossible to write \( M(\text{Re}(\lambda)) = f(\text{Re}(\lambda)) \tilde{M} \) for scalar \( f \) and fixed matrix \( \tilde{M} \).

Also the matrix \( C \) depends on \( \text{Re}(\lambda) \). In order to compute the entries \( c_{i\ell} \), we have to compute the gradient \( \nabla N_\ell \) of the Lagrange basis function \( N_\ell \). We note that \( \nabla N_\ell \) associated with the mesh node \( \ell \) is a linear combination of the Nédélec basis functions associated with edges incident to node \( \ell \). The linear factors are either \(-1\) or \(1\), depending on the direction of the edge. Similarly, the gradient of the quadratic Lagrange basis function associated with an edge mid-point is equal to the quadratic Nédélec basis function associated with that
where the coefficients \( y_{j\ell} \) are either 1, 0, or \(-1\). By inserting (4.8) into the definition of \( C \) and considering the definition of \( M \) in (4.7), it is obvious that the columns of \( C \) are linear combinations of the columns of \( M \), i.e.,

\[
C = M[y_1, \ldots, y_m] = MY,
\]

where the components \( y_{j\ell} \) form a sparse rectangular matrix \( Y \in \mathbb{R}^{n \times m} \) with at most two nonzeros per column. \( Y \) can be easily built from the mesh data.

The desired eigensolutions \((\lambda, x)\) satisfy \( Y^T M x = C^T x = 0 \). This constraint eliminates the zero eigenvalues of \( T(\lambda) \). We use the \( M \)-orthogonal projector \( P \) [23]

\[
P = I - YH^{-1}C^T, \quad H = Y^T MY = Y^T C,
\]

to force the search space into the null space of \( C^T \). This implies that all vectors satisfy the divergence-free condition. The projector \( P \) is applied to a vector as

\[
P x = x - Y(H^{-1}(C^T x)).
\]

The most expensive step is solving with \( H \in \mathbb{C}^{m \times m} \). Note that

\[
h_{ij} = \int_{\Omega} \varepsilon_r \nabla N_i \cdot \nabla N_j \, dx, \quad 1 \leq i, j \leq m,
\]

such that \( H \) depends on \( \text{Re}(\lambda) \). If \( \varepsilon_r \) was constant \( H \) would be a scaled Poisson matrix [4]. We do, however, not construct \( H \) by means of (4.12) but use the formula \( H = C^T Y \) in (4.10).

From (4.8) and the well-known identity \( \nabla \times \nabla N_\ell = 0 \) we obtain that \( AY = O \). In fact, the nullspace of \( A \) coincides with the range of \( Y \). Therefore, enforcing the divergence-free condition entails that the computed eigenvectors are \( M \)-orthogonal to the range of \( Y \). Furthermore, \( A \) is positive definite on the divergence-free subspace [4].

### 4.3 Nonlinear eigensolver

The method of successive linear problems and nonlinear Jacobi–Davidson (JD) methods are popular methods for solving nonlinear eigenvalue problems, see the survey by Voss [81]. The nonlinear Jacobi–Davidson methods have mostly been applied to matrix polynomials or eigenvalue problems of the form
4.3 Nonlinear eigensolver

\[ \sum_{j=1}^{p} f_j(\lambda) K_j x = 0 \] [67, 68]. Following Betcke and Voss [8, 81] we developed our own parallel nonlinear JD solver and successfully solved a nonlinear eigenvalue problem of the form

\[ A x + \lambda^2 R x - \lambda^2 M x = 0, \quad C^T x = 0, \] (4.13)

to analyze electromagnetic cavities using surface impedance boundary conditions [30]. In contrast to (4.13) and the mentioned nonlinear eigenvalue problems, the matrices \( M \) and \( C \) in (4.6) depend on the eigenvalue and have to be rebuilt in each iteration step of the Jacobi–Davidson method. This turned out to be very expensive. Since \( M \) depends on \( \lambda \) in a non-differentiable way, it is not clear how to compute \( \frac{dT}{d\lambda} \) that is required in nonlinear Jacobi–Davidson as well as in the method of successive linear problems.

We therefore resort to a two-parameter formulation of (4.6) that we used similarly and successfully in the 3-dimensional eigenmodal analysis of plasmonic nanostructures [29],

\[ Q(\lambda, \rho) x \equiv A x + \lambda R x - \lambda^2 M(\rho) x = 0, \quad C(\rho)^T x = 0, \]

\[ \rho = \text{Re}(\lambda). \] (4.14)

Since \( T(\lambda) = Q(\lambda, \text{Re}(\lambda)) \), solutions \((\lambda, \text{Re}(\lambda), x)\) of (4.14) are related to solutions \((\lambda, x)\) of (4.6). In a straightforward nonlinear Jacobi–Davidson method for solving (4.14), a projected nonlinear eigenvalue problem

\[ V^T Q(\lambda, \rho) V x = 0 \] (4.15)

has to be solved in every iteration step. Since \( M = M(\rho) \), forming \( V^T Q(\lambda, \rho) V \) is very expensive. Thus, we want to avoid solving (4.15). Note that there is a number of papers on JD methods for nonlinear two-parameter eigenvalue problems, see, e.g., [49, 36], where only constant matrices occur and the projected eigenvalue problem is linear or at least polynomial. The basic idea of our procedure to compute an eigenvalue in the vicinity of a target value \( \tau \) is as follows. Setting \( \rho_0 = \tau \) we solve a sequence of constrained quadratic eigenproblems.

For \( k = 0, 1, \ldots \) until convergence do

Determine the eigenvalue \( \lambda_{k+1} \) closest to \( \tau \) of

\[ Q(\lambda, \rho_k) x = 0, \quad C(\rho_k)^T x = 0. \] (4.16)

Set \( \rho_{k+1} = \text{Re}(\lambda_{k+1}). \)

The quadratic eigenproblem (4.16) is solved by the nonlinear Jacobi–Davidson
4 A nonlinear eigensolver for the analysis of dispersive nanostructures

(NLJD) algorithm. The iteration is considered converged if the condition

$$\left| \rho_{k+1} - \rho_k \right| \leq \varepsilon_1 |\rho_k|$$

is satisfied (\(\varepsilon_1\) is the tolerance) together with the requirement that the NLJD algorithm has converged. We often have to stabilize the update of \(\rho_{k+1}\) by introducing a damping term \(d\),

$$\rho_{k+1} = \rho_k + d \left( \text{Re}(\lambda_{k+1}) - \rho_k \right). \quad (4.17)$$

Algorithm 3 Nonlinear eigensolver FemaxxNano for computing an eigenpair of (4.6). The procedure repeatedly calls the nonlinear JD algorithm 4.

**Input:** Convergence tolerance \(\varepsilon_1\); damping parameter \(0 < d \leq 1\); target \(\tau \in \mathbb{R}\).

**Output:** Eigenpair \((\lambda, u)\) of \(T(\lambda)x = 0\) that is closest to \(\tau\).

1: Set \(\rho_0 = \tau\); \(k = 0\).
2: Construct the constant matrices \(A, R\).
3: Choose an initial search vector \(v_0\), \(\|v_0\|_2 = 1\).
4: while \(k \leq\) the maximally allowed number of iterations do
5: Prepare the matrices \(M(\rho_k), C(\rho_k), H(\rho_k),\) and \(T(\rho_k) = A + \rho_k R - \rho_k^2 M(\rho_k)\) for the call of the NLJD algorithm.
6: Determine a preconditioner \(K \approx T^{-1}\).
7: Invoke the NLJD algorithm 4 to compute the eigenpair \((\lambda_{k+1}, u_{k+1})\) of \(Q(\lambda, \rho_k)x = 0\) that is closest to \(\rho_k\). The algorithm returns \(\text{JD\_converged=TRUE}\) if it converged to the desired accuracy.
8: if \(|\rho_k - \text{Re}(\lambda_{k+1})| < \varepsilon_1 |\rho_k|\) and \(\text{JD\_converged}\) then
9: Return \((\lambda_{k+1}, u_{k+1})\) as the eigenpair.
10: end if
11: \(\rho_{k+1} = \rho_k + d \left( \text{Re}(\lambda_{k+1}) - \rho_k \right);\) \(v_0 = u_{k+1}\);
12: end while

Details of the procedure are given in Algorithm 3 which forms the principle loop of our FemaxxNano eigensolver. Since we are interested in those eigenvalues that are close to the real axis, we choose the target \(\tau\) to be real. Note that electromagnetic modes with large \(\text{Im}(\lambda)\) decay fast, such that they are hardly excited or observed in reality.

In step 3 of Algorithm 3, we construct the initial vector \(v_0\) such that the associated finite element function is 1 inside \(\Omega_1\) (see Fig. 4.1) and 0 outside. The solver often failed to converge with random initial vectors. In step 5
the matrix $T(\rho_k)$ is generated for convenience. It is used to construct the preconditioner in the next step. We discuss in Section 4.4 how to choose the damping parameter $d$.

We terminate the iteration if the change in the real parts of the eigenvalues are sufficiently small, provided that the inner NLJD iteration also finished satisfactory, i.e., the residual $\|Q(\lambda_{k+1}, \rho_k)u_{k+1}\|_2$ is small.

**Algorithm 4** Nonlinear Jacobi–Davidson algorithm for computing a solution of the quadratic eigenvalue problem (4.16).

**Input:** Matrices $A, R, M, T, Y, H, C, K$; initial vector $v_0$; target $\rho_k$; decay parameter $\gamma$; $j^{\text{max}}$; $\epsilon_2^{\text{real}}, \epsilon_2^{\text{imag}}$.

**Output:** Eigenpair $(\lambda, u)$ of $Q(\lambda, \rho_k)x = 0$ that is closest to $\rho_k$; convergence indicator $\text{JD}_{\text{converged}}$.

if $\text{JD}_{\text{converged}}$ = TRUE: eigenpair $(\lambda, u)$ satisfies certain convergence criteria.

1: Set initial search space $V = \{v_0\}$; $j = 0$; $\text{JD}_{\text{converged}}$ = FALSE.
2: Project $v_0$ on $\mathcal{R}(C)^\perp$, i.e., $v_0 = (I - YH^{-1}C^T)v_0$.
3: while true do
   4:   $j = j + 1$.
   5:   Compute the eigenpair $(\lambda_j, x)$ of the projected eigenproblem $(V^*AV + \lambda V^*RV - \lambda^2 V^*MV)x = 0$ with $\lambda_j$ closest to $\rho_k$.
   6:   Set Ritz vector $u = Vx$, $u = u/\|u\|$; residual $r = (A + \lambda_j R - \lambda_j^2 M)u$.
   7:   if $|\text{Re}(\lambda_j - \lambda_{j-1})| < \epsilon_2^{\text{real}}|\text{Re}(\lambda_{j-1})|$ and $|\text{Im}(\lambda_j - \lambda_{j-1})| < \epsilon_2^{\text{imag}}|\text{Im}(\lambda_{j-1})|$ then
      8:      Set convergence flag $\text{JD}_{\text{converged}}$ = TRUE.
   9:   end if
10:  if $\text{JD}_{\text{converged}}$ or $j > j^{\text{max}}$ then
11:     return the Ritz pair $(\lambda_j, u)$ and $\text{JD}_{\text{converged}}$.
12:  end if
13:  Set $p = Ru - 2\lambda_j Mu$.
14:  Find an approximate solution of the correction equation
   \[
   \left( I - \frac{pu^*}{u^*p} \right) T \left( I - \frac{uu^*}{u^*u} \right) t = -r, \quad t \perp u. \tag{4.18}
   \]
   by a Krylov solver preconditioned by $K$.
15:  Project $t$ on $\mathcal{R}(C)^\perp$: $t = (I - YH^{-1}C^T)t$.
16:  Orthogonalize $t$ against the current search space, $t = t - VV^*t$, and normalize $t = t/\|t\|_2$.
17:  Expand subspace $V = [V, t]$.
18:  end while

The NLJD algorithm 4 is taken from Voss [81] with only a few modifications. Algorithm 4 contains the steps that apply the projector $P$, cf. steps 2 and 15. Another type of convergence criterion is implemented, too, cf. step 7. The details of the algorithm are as follows.

(a) We enforce the divergence-free condition in step 2 and 15, by applying the
projector $P$ in (4.10)–(4.11). We solve the systems that involve the matrix $H$ by the stabilized Bi-Conjugate Gradient (Bi-CGstab) method proposed by van der Vorst [76]. We found Bi-CGstab a good compromise between COCG and GMRES. The complex-orthogonal conjugate gradient method COCG is a variant of Bi-CG that exploits the complex-symmetry of the underlying system [77, 5]. It is not as stable as the Generalized Minimal Residual (GMRES) method [65] which however has high computational complexity and memory requirement. We found Jacobi preconditioning very efficient for solving with the matrix $H$.

(b) The projected eigenproblem $(V^*AV + \lambda V^*RV - \lambda^2 V^*MV)x = 0$ in step 5 is a small-sized quadratic eigenvalue problem. We solve it with the QZ algorithm after linearization [29, 74].

(c) We are interested in eigenvalues that are close to the real axis (i.e., $\text{Re}(\lambda) > |\text{Im}(\lambda)|$), so $\text{Im}(\lambda)$ could be inaccurate even if $\text{Re}(\lambda)$ has converged. Therefore the stopping criterion in step 7 checks the real and imaginary parts of $\lambda_j - \lambda_{j-1}$ individually, with two tolerances $\varepsilon_2^{\text{real}}$ and $\varepsilon_2^{\text{imag}}$. If the stopping criterion is satisfied, the normalized residual $\|r\|_2$ is also small provided that $T'(\lambda)^{-1}$ is bounded in the vicinity of an eigenvalue.

(d) In step 10, NLJD may terminate without satisfying the convergence criteria if the search space has reached the maximally admitted dimension $j^{\text{max}}$. In this situation $\text{JD\_converged}$ is FALSE. Note that there is no restart in our NLJD solver.

(e) The correction equation (4.18) in step 14 of Algorithm 4 is again solved by the Bi-CGstab algorithm. Let the maximum number of iteration steps for BI-CGStab (inner iteration) be $it^{\text{max}}_{\text{corr}}$. We discuss the preconditioner $K \approx T^{-1}$ in Section 5. It is not necessary to solve the correction equation accurately. So, $it^{\text{max}}_{\text{corr}}$ need not be very large. On the other hand, setting $it^{\text{max}}_{\text{corr}}$ very small can be dangerous as then the computed corrections may be so ineffective that the NLJD solver fails to converge.

(f) Particularly in the initial NLJD iterations, the computed Ritz value is often far from the desired eigenvalue. Consequently, the calculated correction $t$ in step 14 is generally ineffective, even if the correction equation is solved very accurately [23]. In such situations, it makes sense to terminate the Bi-CGstab iterations as early as possible.

For this reason, we use a variable stopping criterion for Bi-CGstab,

$$\|\tilde{r}_i\|_2 < \gamma^{-j_\text{step}}\|\tilde{r}_0\|_2,$$

(4.19)

as suggested by Fokkema et al. [19]. Here, $\tilde{r}_i$ is the residual associated with
4.4 Numerical experiments

4.4.1 Numerical experiments

We first recalculated with FemaxxNano the nanostructures analyzed in [29], i.e., nanocuboids, nanospheres and optical dipole antennas. In those cases, the permittivity $\varepsilon_r$ was considered non-dispersive. The numerical results agreed with those in [29]. Therefore, we consider the algorithm validated. In this paper, we analyze a gold nanoantenna [70]. In Fig. 4.2, we plot the dielectric permittivity of gold in the spectral range of $[0.64\text{ eV}, 6.60\text{ eV}]$ [40]. The real and the imaginary parts of $\varepsilon_r$ are displayed separately. In our numerical experiments, we study the eigenmodes within the infrared spectral range of $S_{if} = [0.64\text{ eV}, 1.46\text{ eV}]$ or, equivalently, $[812\text{ nm}, 1937\text{ nm}]$. Gold obviously exhibits significant dispersion in $S_{if}$, where the module of $D (= \omega \frac{d\varepsilon_r}{d\omega})$ is above 50, and grows even larger than 300, see Fig. (4.3).

4.4.2 Numerical experiments

We first recalculated with FemaxxNano the nanostructures analyzed in [29], i.e., nanocuboids, nanospheres and optical dipole antennas. In those cases, the permittivity $\varepsilon_r$ was considered non-dispersive. The numerical results agreed with those in [29]. Therefore, we consider the algorithm validated. In this paper, we analyze a gold nanoantenna [70]. In Fig. 4.2, we plot the dielectric permittivity of gold in the spectral range of $[0.64\text{ eV}, 6.60\text{ eV}]$ [40]. The real and the imaginary parts of $\varepsilon_r$ are displayed separately. In our numerical experiments, we study the eigenmodes within the infrared spectral range of $S_{if} = [0.64\text{ eV}, 1.46\text{ eV}]$ or, equivalently, $[812\text{ nm}, 1937\text{ nm}]$. Gold obviously exhibits significant dispersion in $S_{if}$, where the module of $D (= \omega \frac{d\varepsilon_r}{d\omega})$ is above 50, and grows even larger than 300, see Fig. (4.3).
4 A nonlinear eigensolver for the analysis of dispersive nanostructures

Figure 4.2: The relative permittivity $\varepsilon_r$ for gold as a function of the photon energy [40]

Figure 4.3: Plot of the parameter $D$. 
According to our experience, a large $D$ requires a small damping parameter $d$ in (4.17) to stabilize the update of $\rho_k$. Therefore, in this Chapter we chose $d = 0.1$ in the highly dispersive range $S_{if}$. If we choose a larger $d$, say $d = 0.2$, then we observed faster convergence to some of the electromagnetic eigenmodes in $S_{if}$, but no convergence at all to others. In [29], the nanostructures (nanocuboids and nanospheres) have plasmonic resonances only in visible or the ultraviolet spectral range [2.8 eV, 6.6 eV], where the module of $D$ is much smaller compared to that in $S_{if}$. Then we can safely set $d$ between 0.3 and 0.5. For a dielectric resonator antenna (DRA) [29], the resonant frequency is in the microwave region of the electromagnetic spectrum, where the dielectric dispersion is almost negligible, i.e. $D \approx 0$. Then we can safely choose $d$ between 0.9 and 1. So, the appropriate selection of the damping parameter $d$ strongly depends on the material dispersion in the considered electromagnetic spectral range. If the associated effects are high, e.g. gold in the infrared spectral range, then the nonlinearity of the matrices $M$ or $C$ are strong and $d$ should be chosen small in order to reduce the chance of missing certain eigenmodes.

The sketch of the structure is shown in Fig. 4.4. The antenna has two disks, each being a cylinder with a cap of semi-ellipsoidal shape. The height of each disk is $h = 30$ nm, and the radius is $r$ that we will set 60 nm and 120 nm, respectively, in our experiments. The semi-axes of the cap are $r$ along the $x$ and $y$ axes and $w = 3$ nm along the $z$ axis. The gap between the two disks is 1 nm. The radius of the spherical computational domain is $R = 400$ nm.

For all simulations, the three tolerances $\epsilon_1 = \epsilon_2^{\text{real}} = 0.5\%$ and $\epsilon_2^{\text{imag}} = 1\%$ are used in the stopping criteria. The decay parameter in (4.19) is $\gamma = 1.2$. All simulations were carried out on the Cray XT6 of the Swiss National Supercomputing Centre (CSCS) [14].

### 4.4.1 A ‘small’ nanoantenna

We now consider a ‘small’ nanoantenna with $r = 60$ nm. First, we study the solver behavior when linear elements are used. We consider two meshes. The coarse mesh consists of 71’003 tetrahedra. The finite element discretization counts 88’021 degrees of the freedom (dof) for $T$ and 13’823 dof for $H$. The fine mesh contains 436’691 tetrahedra. $T$ and $H$ have orders 531’355 and 81’515, respectively. We set $j^{\text{max}} = 15$ and $it_{\text{corr}}^{\text{max}} = 30$, respectively.

The preconditioner for the correction equation (4.18) is Jacobi, i.e., $K = \text{diag}(T)$. At present, we have not implemented advanced complex-valued preconditioners such as AMG preconditioners [3], but the convergence (as discussed later) is still quite satisfactory. In particular, in a parallel environment, Jacobi preconditioning is easy to implement and very memory-efficient.

We create a set of four targets, $S = \{\tau_0 = 0.71$ eV, $\tau_1 = 0.91$ eV, $\tau_2 =$
Figure 4.4: (a) Cross sectional view and (b) a 3-D sketch of the nanoantenna with cylindrical disks.

1.11 eV, \( \tau_3 = 1.31 \text{eV} \), with a spacing of \( \Delta \tau = 0.2 \text{eV} \). \( S \) is contained in the spectral range of interest \( S_{if} \).

We obtain 3 eigenmodes with \( \| r \|_2 \) between \( 10^{-5} \) to \( 5 \times 10^{-5} \). We denote the 3 eigenmodes M1, M2 and M3, respectively. In Tables 4.1 and 4.2 we present three relevant quantities that are derived from the eigenvalues \( \lambda \). The photon energy of the resonant eigenmode \( E_{\text{res}} \), the wavelength at the resonant eigenmode \( \lambda_{\text{res}} \), and the quality factor \( Q \), which are formally defined as

\[
E_{\text{res}} = \hbar c \text{Re}(\lambda), \quad \lambda_{\text{res}} = \frac{2\pi c \hbar}{E_{\text{res}}}, \quad Q = \frac{\text{Re}(\lambda)}{2\text{Im}(\lambda)},
\]

where \( \hbar \) is the Planck constant and \( c \) is the speed of light in vacuum.

Table 4.1: Numerical results for the small nanoantenna (\( r = 60 \text{nm} \)) and a fine mesh with 436,691 tetrahedra.

<table>
<thead>
<tr>
<th>mode</th>
<th>linear elements (dof = 531,355)</th>
<th>quadratic elements (dof = 2,835,774)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( E_{\text{res}} ) (eV)</td>
<td>( \lambda_{\text{res}} ) (nm)</td>
</tr>
<tr>
<td>M1</td>
<td>0.761</td>
<td>1629</td>
</tr>
<tr>
<td>M2</td>
<td>1.088</td>
<td>1139</td>
</tr>
<tr>
<td>M3</td>
<td>1.255</td>
<td>988</td>
</tr>
</tbody>
</table>

The results for the fine mesh in Table 4.1 have been obtained on 64 cores.
Table 4.2: Numeric results for the small nanoantenna \((r = 60 \text{ nm})\) and a coarse mesh with 71’003 tetrahedra.

<table>
<thead>
<tr>
<th>mode</th>
<th>linear elements (dof = 88’021)</th>
<th>quadratic elements (dof = 466’446)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(E_{\text{res}}) (eV)</td>
<td>(\lambda_{\text{res}}) (nm)</td>
</tr>
<tr>
<td>M1</td>
<td>0.752</td>
<td>1648</td>
</tr>
<tr>
<td>M2</td>
<td>1.071</td>
<td>1157</td>
</tr>
<tr>
<td>M3</td>
<td>1.257</td>
<td>986</td>
</tr>
</tbody>
</table>

of the Cray XT6. The total computational time was below 3 minutes. The results for the coarse mesh in Table 4.2 took about 2 minutes on 8 cores of the Cray XT6.

The convergence history for linear elements with the fine mesh is shown in Fig. 4.5(a). In total 106 NLJD iteration steps, FemaxxNano computes 3 eigen-modes. For both targets \(\tau_0 = 0.71\) eV and \(\tau_1 = 0.91\) eV, FemaxxNano converges to the mode M1 (0.761 eV). The mode M2 (1.088 eV) is found near \(\tau_2 = 1.11\) eV, and the mode M3 (1.255 eV) is found near \(\tau_3 = 1.31\) eV. We note that the convergence may not be steady. The noticeable rises of \(\|r\|_2\) usually happens in the outer iteration steps, when \(M, T, C\) and \(H\) are recomputed with an updated \(\rho_{k1}\) (step 5 in Alg. 3). Vectors \(x_k\) with \(\|x_k\| = 1\) that gave rise to small residuals \(Q(\lambda_k, \rho_k)x_k\) lead to larger residuals \(Q(\lambda_k, \rho_{k+1})x_k\) at the start of the next NLJD iteration.

For each NLJD iteration, the average number of BI-CGStab iterations is 18.1 for approximately solving Eq. (4.18). In 42 out of totally 106 NLJD steps, BI-CGStab returns with a non-converged correction \(t\) after \(i_{\text{corr}}^{\text{max}} = 30\) iterations.

Now, we turn to the quadratic elements. In both discretizations the number of tetrahedra is unchanged. In the coarse mesh we now have 466’446 dof for \(T\) and 101’844 dof for \(H\). In the fine mesh the finite element discretization leads to 2’835’774 dof for \(T\) and 612’870 dof for \(H\). We set \(i_{\text{corr}}^{\text{max}} = 60\) and \(j^{\text{max}} = 25\) and use the same sampling as before.

In Femaxx, we use hierarchical bases for both the Nédélec and the Lagrange basis functions. Therefore, by numbering the linear before quadratic degrees of freedom, the matrices \(A\), \(M\), \(T\), and \(H\) all have a 2-by-2 block structure. It is natural to use a hierarchical basis preconditioner [3, 23] for solving the correction equation. The crucial part of the correction equation is solving for \(z\) in

\[
Tz = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix},
\]
Figure 4.5: The convergence history for the fine mesh. The $y$ axis denotes the normalized residual $\|r\|_2$, and the $x$ axis denotes the NLJD iteration step (IT). Linear (a) and quadratic (b) elements are used.
4.4 Numerical experiments

where $T_{11}$ corresponds to the bilinear forms involving linear basis functions. We replace solving with $T$ by of one step of symmetric block Gauss–Seidel iteration \[4],

$$
\begin{align*}
    z_1 &= T_{11}^{-1} f_1, \\
    z_2 &= T_{22}^{-1} (f_2 - T_{21} z_1), \\
    z_1 &= T_{11}^{-1} (f_1 - T_{12} z_2).
\end{align*}
$$

In fact, we even approximate the diagonal blocks $T_{11}$ and $T_{22}$ by their diagonals – which we call Jacobi preconditioning.

Using 256 cores of the Cray XT6 for the fine mesh, the total NLJD iteration steps is 200, and the total computational time is below 25 minutes. Again, we obtain the 3 eigenmodes M1, M2, and M3 with $\|r\|_2$ between $10^{-5}$ to $5 \times 10^{-5}$. For each NLJD iteration, the average number of BI-CGStab iterations is 47.5 for approximately solving (4.18). In 114 out of totally 200 NLJD steps, BI-CGStab returns with a non-converged correction $t$ after $it_{\text{corr}}^{\text{max}} = 60$ iterations. The resonances and the corresponding Q factors are also given in Tables 4.1 and 4.2. The results for quadratic elements on the coarse mesh took about 4 minutes on 64 cores of the Cray XT6, which is only about 30% more than that for linear elements on the fine mesh.

The convergence history for the quadratic elements on the fine mesh is plotted in Fig. 4.5(b). The iteration count increases considerably when compared to linear elements. In particular the iterations starting from targets $\tau_0$ and $\tau_2$ converge in quite an erratic way.

The numerical results for the quadratic elements on the fine mesh are the most accurate, such that we use them as the reference solutions. Then, the quadratic elements on the coarse mesh give more accurate results (below 0.5% for $E_{\text{res}}$ and 1% for $Q$) than the linear elements on the fine mesh (below 2% for $E_{\text{res}}$ and 4% for $Q$). The results with the linear elements on the coarse mesh differ from those with the quadratic elements on the fine mesh by up to about 3.5% for $E_{\text{res}}$ and 10% for $Q$.

If we refine the sampling with $\tau_0 = 0.66$ eV and $\Delta \tau = 0.1$ eV, then there are 8 sampling points in $S_d$. We find no other eigenmodes than M1, M2 and M3. Therefore, we assume that we did not miss an eigenpair in the investigated range.

The electric field distributions $|E|$ obtained on the fine mesh are shown in Figs. 4.7(a) to 4.7(d). We observe drastic field enhancement in the gap between the gold disks, while only scant energy is radiated to the far field.
4 A nonlinear eigensolver for the analysis of dispersive nanostructures

4.4.2 A ‘large’ nanoantenna

We now investigate a ‘large’ nanoantenna with radius $r = 120$ nm. The other geometrical parameters are the same as those of the small nanoantenna. We use a mesh with 380'316 tetrahedra and quadratic elements are used. The discretization counts 2'478'574 dof for $T$ and 537'545 for $H$. We set $i_{\text{corr}}^{\text{max}} = 60$, and $j^{\text{max}} = 25$. Again, we use the hierarchical basis preconditioner.

Table 4.3: Numerical analysis of a gold nanoantenna with $r = 120$ nm. Here the quadratic elements are employed.

<table>
<thead>
<tr>
<th>mode</th>
<th>$E_{\text{res}}$ (eV)</th>
<th>$\lambda_{\text{res}}$ (nm)</th>
<th>$Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M2</td>
<td>M3</td>
<td>M4</td>
</tr>
<tr>
<td>$E_{\text{res}}$ (eV)</td>
<td>0.674</td>
<td>0.792</td>
<td>0.860</td>
</tr>
<tr>
<td>$\lambda_{\text{res}}$ (nm)</td>
<td>1839</td>
<td>1565</td>
<td>1441</td>
</tr>
<tr>
<td>$Q$</td>
<td>2.90</td>
<td>3.87</td>
<td>2.98</td>
</tr>
</tbody>
</table>

Figure 4.6: Distribution of the eigenmodes for the gold nanoantenna within the infrared range $S_{\text{if}}$; (+) radius $r = 60$ nm; (●) radius $r = 120$ nm. The results are obtained by quadratic elements.

This structure supports many more eigenmodes, and some of them are quite close. Therefore we need a higher resolution sampling of targets. We chose $\Delta \tau = 0.05$ eV. The now 17 targets are $\tau_0 = 0.65$ eV, $\tau_1 = 0.70$ eV, up to $\tau_{16} = 1.45$ eV. They are again contained in $S_{\text{if}}$.
4.4 Numerical experiments

Figure 4.7: The electric field distribution $|\mathbf{E}|$ in the gap of gold nanoantennas obtained with quadratic elements. (a)–(d) radius $r = 60 \text{ nm}$; (e)–(k) radius $r = 120 \text{ nm}$. (a) is visualized on the $yz$-plane, and all the other sub-figures are visualized on the $xy$-plane through the center of the gap.
On 256 cores we computed 10 eigenmodes in slightly less than 2 hours. The eigenmodes obtained are given in Table 4.3. The normalized residuals $\|r\|_2$ were between $10^{-5}$ to $5 \times 10^{-4}$. The higher order modes often present relatively lower numerical accuracy.

The first two modes ($E_{\text{res}} = 0.674 \text{ eV}$ and $0.792 \text{ eV}$) have similar electric field distributions as the first two of the small nanoantenna in Figs. 4.7(c) and 4.7(d), so we still call them M2 and M3. Note that we did not find a mode of shape M1 in the range $S_{1f}$. Such a mode may exist in the spectral range below $0.64 \text{ eV}$; but there the experimental data for $\varepsilon_r$ is not available [40].

The next 8 modes are called M4 to M11, with electric field distributions shown in Figs. 4.7(e) to 4.7(l), respectively. Interestingly, $E_{\text{res}}$ for M8 is only 1.9% larger than that of M7, so they are spectrally close. However, their field patterns are very different.

Again, by refining $S$ to 33 sampling points with $\Delta \tau = 0.025 \text{ eV}$ such that $\tau_0 = 0.65 \text{ eV}$ and $\tau_{32} = 1.45 \text{ eV}$, we find no other eigenmodes than M2 to M11. Therefore, it is unlikely that we missed an eigenmode in the investigated interval close to the real axis.

We summarize the resonances and $Q$ factors for all computed eigenmodes in Fig. 4.6.

4.5 Discussions and Conclusions

We have implemented the electromagnetic eigenvalue solver code FemaxxNano for the analysis of plasmonic nanostructures that are mainly characterized by metals with highly dispersive dielectric permittivity in the optical region of the electromagnetic spectrum. Via the discretization of the time-harmonic electric field vector wave, or curl-curl, equation with the 3-dimensional finite element method (FEM) a constrained, nonlinear eigenvalue problem, Eq. (4.6), is built-up. The algorithm of our eigensolver consists of two loops. In the outer loop, a sequence of constrained quadratic eigenvalue problems (3.15) is constructed. These quadratic eigenvalue problems are solved in the inner loop with the nonlinear Jacobi-Davidson (NLJD) method. We find that the selection of an appropriate initial vector for FemaxxNano algorithm is essential for convergence. Usually, the eigensolver fails to converge with random initial vectors. Therefore, we construct the initial vector such that the associated finite element function is 1 inside the nanostructure $\Omega_1$ and 0 outside. Also, in order to improve the robustness of FemaxxNano, the so-called damping parameter must be carefully chosen according to the associated material dispersion properties. In order to calculate multiple eigenpairs, we create a set of targets and search for the corresponding eigenpair in the vicinity of each target. Our
solver FemaxxNano has been parallelized for distributed memory computers. We show results that have been obtained using 8 up to 256 cores. The numbers show excellent accuracy at relatively low computational cost, compared to our previous approach [29].

We have applied the eigensolver to gold optical antenna structures [69]. They support eigenmodes in the highly dispersive infrared spectral range. Detailed numerical results and eigensolver behavior are presented and discussed. Both linear and quadratic Nédélec elements are used for comparison. We find that quadratic elements provide relatively more accurate solutions, while the corresponding iterations converge in quite an erratic way. By appropriately refining the targets sampling set, even close eigenmodes are calculated successfully. The numerical experiments demonstrate the capability of our eigensolver for modeling complicated material dispersion properties as found in gold.

We hope that FemaxxNano will become a useful numerical simulation tool to further the understanding of dispersive plasmonic nanostructures.
Electromagnetic eigenmodal analysis of a molecular nanosensor

Here, we employ the 3-dimensional finite element eigensolver tool, Femaxx, to theoretically analyze the electromagnetic eigenmodes of a nanometer sized molecular sensor device. The nanosensor device is based on the optical dipole antenna configuration [54, 9, 53], which consists of two metallic dipoles, separated by a small gap, situated on a substrate. We compute the resonant frequencies, the sensitivity and the figure of merit (FOM), and access the sensor performance.

5.1 Introduction

In a previous publication [29] we have introduced the Femaxx numerical code for solving the electromagnetic eigenmodes of nanometer sized optical structures, by 3-dimensional finite element analysis. Several examples are presented in [29]. One of them, the optical dipole antenna, represents an important, elementary configuration in many detection and sensing devices. For details of our numerical method and the basis of the optical dipole antenna, we refer the readers to [29].

Here, we use Femaxx to study the physical behavior of the optical dipole antenna when employed as a molecular nanosensor [47, 38, 1]. According to the results in [29] the optical dipole antenna possesses two electromagnetic modes in the optical region of the electromagnetic spectrum, the bright mode and the dark mode. The bright mode radiates non-negligible energy to the far field zone, with a drastic field enhancement arising in the gap [29]. On the other hand, the dark mode localizes almost all energy in the vicinity of the antenna surface, with only scant energy radiated away. We investigate the sensing capability for these two different modes.
5 Electromagnetic eigenmodal analysis of a molecular nanosensor

5.2 Numerical experiment

We show the geometry and sketch of the plasmonic nanosensor in Figs. 5.1(a) and 5.1(b), respectively. The antenna has two gold arms, each of which has dimensions $a = b = 40$ nm, $l = 100$ nm. In order to better approximate reality, we model the corners of each arm to be rounded, with the radius of curvature being 5 nm. The gap $g$ between the two arms is 10 nm wide. The antenna resides on a silica substrate with relative dielectric permittivity $\epsilon_r = 2.25$, and is surrounded by air with $\epsilon_r = 1$. For the gold arms, we use experimental data for the dispersive complex relative permittivity $\epsilon_r$ taken from [40].

We use the 1st order absorbing boundary condition (ABC) in order to truncate the computational domain and, thus, to keep the tetrahedral mesh size finite. The radius $R_b$ of the spherical computational domain is 400 nm. The optical dipole antenna is completely surrounded by a dielectric molecular layer, with a finite thickness $t$ and the refractive index $n$. Note that the layer also presents rounded corners with a radius of $= 5$ nm, and the gap is also filled by the molecules.

![Diagram of the dipole antenna](image_url)

Figure 5.1: (a) Geometry of the dipole antenna. (b) The sketch for molecular nanosensing: both the top view (left) and the cross section (right) are shown. The yellow area is the optical dipole antenna, and it resides on a silica substrate (gray area) and it is surrounded by a molecule layer (light gray area).

We use a mesh with about 400'000 tetrahedra. The triangulation has to respect all interfaces in the sketch 5.1(b). To reliable resolve the bright mode, we use quadratic Nédélec element basis functions which leads to roughly 2.6 million degrees of freedom (dof) in the discretized, 3-dimensional problem. All
5.2 Numerical experiment

Simulations are performed on Cray XT6 at Swiss National Supercomputing Centre [14], with 200 to 300 cores employed.

The nanosensor acts as a transducer so that a relatively small change in $n$ is converted into a spectral red-shift of the resonance $\lambda_{\text{res}}$ [1]. We evaluate the sensitivity of the nanosensor based on this shift. Let $t = 5$ nm. We vary $n$ s.t. $n = 1.0, 1.2, 1.3, 1.4, 1.5$ and $1.6$, respectively. (Then, the corresponding dielectric constants are $\varepsilon_r = n^2$.) In Fig. 5.2 we plot the linear relationship between the resonance wavelength $\lambda_{\text{res}}$ and the refractive index $n$ for the bright mode. The regression equation evaluates to $\lambda_{\text{res}} = 172n + 604$ (nm). Therefore the sensitivity is 172 nm per unit of refractive index (RIU). The bright mode shows high sensitivity in the near-infrared region, where the metallic dispersion is considerably strong [40].

![Figure 5.2: The resonance $\lambda_{\text{res}}$ of the bright mode vs. the refractive index $n$ of the molecule layer. The thickness of the layer is $t = 5$ nm.](image)

We also study the sensitivity as a function of the thickness $t$ of the molecule layer. We vary $t$ s.t. $t = 2, 5, 10, 20$ and $40$ (nm), respectively. The corresponding sensitivity for each $t$ is plotted in Fig. 5.3(a). When increasing $t$ the sensitivity of the nanosensor improves. For instance, if $t = 40$ nm, the sensitivity is almost 300 nm/RIU. We comment that the increase of the sensitivity tends to flatten out as it was observed and discussed in [47]. We show the electric field distribution in Fig. 5.3. The field is symmetric with respect to the $xz$–plane and $yz$–plane, Due to the substrate and molecular layer effects, the field is not symmetric with respect to the $xy$–plane.

If $t$ is so large that the whole vacuum area in Fig. 5.1(b) is filled with molecules, the sensitivity reaches up to 404 nm/RIU or 0.66 eV/RIU, which is higher than many commonly used nanosensors [47].
Figure 5.3: (a) The sensitivity (nm/RIU) of the nanosensor vs. the thickness of the molecule layer $t$ (nm). (b) Electric field distribution $|\mathbf{E}|$ in the vicinity of the antenna surface. The layer has $n = 1.5$ and $t = 40\, \text{nm}$. The resonant wavelength is $\lambda_{res} = 924\, \text{nm}$. (b) is evaluated on the $xy$-plane through the center of the antenna arms.
5.3 Summary

If there is no molecule layer, i.e. \( t=0 \), then the bright mode resonance is \( E_{res} = 1.60 \text{ eV} \), i.e., 775 nm, and the corresponding quality factor \( Q \approx 17 \) \cite{29}. The full width at half maximum (FWHM) is computed as \( \text{FWHM} \approx E_{res}/Q \approx 0.094 \text{ eV} \). The figure of merit (FOM) of the nanosensor (defined as sensitivity divided by FWHM) \cite{38} can be evaluated as \( \text{FOM} = 0.66/0.094 \approx 7.2 \), which is higher than nanoparticle sensors previously reported in \cite{47, 38}.

Finally, we comment that the dark mode, that radiates scant energy into the far field zone \cite{29}, is much less favorable for sensing. According to our numerical results, when \( t = 5 \text{ nm} \) the dark mode sensitivity is only 30 nm/RIU.

5.3 Summary

We have presented the 3-dimensional electromagnetic eigenmodal analysis of a molecular nanosensor, based on the optical dipole antenna configuration. For the bright mode of the antenna, the sensitivity reaches up to 404 nm/RIU and the corresponding FOM is about 7.2. This chapter demonstrates that Femaxx proves to be flexible and reliable when analyzing realistic plasmonic nanostructures.
Cavity eigenmodal analysis with surface impedance boundary conditions

The material in this chapter was presented at the 11th International Computational Accelerator Physics Conference (ICAP) [30]. The new X-ray Free Electron Laser (SwissFEL) at the Paul Scherrer Institute (PSI) employs, among many other radio frequency elements, a transverse deflecting cavity for beam diagnostics. Since the fabrication process is expensive, an accurate 3-D eigenmodal analysis is indispensable. The software package Femaxx has been developed for solving large scale eigenvalue problems on distributed memory parallel computers. Usually, it is sufficient to assume that the tangential electric field vanishes on the cavity wall (PEC boundary conditions). Of course, in reality, the cavity wall is conductive such that the tangential electrical field on the wall is nonzero. In order to more realistically model the electric field we impose surface impedance boundary conditions (SIBC) arising from the skin effect model. The resulting nonlinear eigenvalue problem is solved with a nonlinear Jacobi–Davidson method. We demonstrate the performance of the method. First, we investigate the fundamental mode of a pillbox cavity. We study resonance, skin depth and quality factor as a function of the cavity wall conductivity. Second, we analyze the transverse deflecting cavity of the SwissFEL to assess the capability of the method for technologically relevant problems.

6.1 Formulation of the problem

We wish to calculate the resonant frequencies and the corresponding field distribution in a dielectric electromagnetic cavity. The cavity wall $\Gamma$ is assumed to be of arbitrary shape; there is no aperture or hole in $\Gamma$. The surface conductivity $\sigma_s$ of $\Gamma$ is large but finite. The interior $\Omega$ of the cavity is assumed to be source-free, and is characterized by $(\mu_0 \mu_r, \varepsilon_0 \varepsilon_r)$. $\mu_0$ and $\varepsilon_0$ are the magnetic permeability and electric permittivity in free space. $\mu_r$ and $\varepsilon_r$ are relative magnetic permeability and relative electric permittivity, respectively.
At microwave frequencies, \( \mu_r \) and \( \varepsilon_r \) can be assumed to be non-dispersive.

In the time-harmonic regime, after eliminating the electric field \( E(x) \), the magnetic field \( H(x) \) satisfies

\[
\nabla \times (\varepsilon^{-1}_{\text{r}} \nabla \times H(x)) - k_0^2 \mu_r H(x) = 0, \quad x \in \Omega,
\]
\[
\nabla \cdot (\mu_r H(x)) = 0, \quad x \in \Omega.
\]

Here, \( k_0 = \tilde{\omega} \sqrt{\mu_0 \varepsilon_0} \) is the complex wave number in free space, \( \tilde{\omega} = \omega + i\alpha \) is the complex angular frequency with \( \omega \) the angular frequency and \( \alpha \) the exponential decay rate. We use the surface impedance boundary condition (SIBC) on \( \Gamma \) [39]

\[
n \times (n \times E(x)) = Z_s n \times H(x), \quad x \in \Gamma.
\]

Here, \( Z_s \) is the complex surface impedance and \( n \) the surface normal vector pointing outwards.

We employ \( Z_s \) based on the theoretical skin effect model [63]

\[
Z_s = \frac{1 + i}{\sigma_s \delta},
\]

where \( \sigma_s \) is the surface conductivity, and \( \delta \) is the skin depth. The real part of \( Z_s \) is the surface resistivity, i.e.,

\[
R_s = \text{Re}(Z_s) = \frac{1}{\sigma_s \delta}.
\]

The skin depth \( \delta \) is [63]

\[
\delta = \sqrt{\frac{2}{\omega \mu_0 \mu_r \sigma_s}}.
\]

\( \delta \) depends on the angular frequency \( \omega \). Note that the skin effect model is appropriate only if \( \sigma_s \) is large enough such that (according to [63]): (1) the conduction current is given by Ohm’s law and the net charge density is zero; (2) the displacement current is negligible in comparison with the current, i.e., \( \omega \varepsilon_r \varepsilon_0 \ll \sigma_s \). With the above two assumptions, we consider the conductor is good, and the loss of the cavity is small. In other words, the decay rate \( \alpha \ll \omega \), and thus \( \omega \approx \tilde{\omega} = k_0 c \), implying that

\[
\delta \approx \frac{2}{k_0 c \mu_0 \mu_r \sigma_s}.
\]

The finite element method (FEM) is a suitable method for arbitrary geometrical scales. In order to apply the FEM we use the weak form of Eq (6.1),

\[
(6.1)
\]
6.1 Formulation of the problem

see [82],

Find $k_0 \in \mathbb{C}$ and $\mathbf{H} \in V$, $\mathbf{H} \neq \mathbf{0}$, such that for all $f \in V$ and all $q \in W$

$$
\int_\Omega \left[ \frac{1}{\varepsilon_r} \nabla \times \mathbf{H} \cdot \nabla \times f - k_0^2 \mu_r \mathbf{H} \cdot f \right] dx \\
+ i k_0 \frac{1}{Z_0} \int_\Gamma (\mathbf{n} \times \mathbf{E}) \cdot f ds = 0, \\
\int_\Omega \mu_r \mathbf{H} \cdot \nabla q dx = 0.
$$

(6.7)

Here, $V$ denotes the functions in $H(\text{curl}; \Omega)$ that satisfy the SIBC boundary conditions and $W = H^1_0(\Omega)$ [39]. $Z_0 = \sqrt{\mu_0/\varepsilon_0}$ is the characteristic impedance of free space.

Plugging in the SIBC (6.2), and using (6.3) and (6.6), we get

$$
i k_0 \frac{1}{Z_0} \int_\Gamma (\mathbf{n} \times \mathbf{E}) \cdot f ds \\
= i k_0 \frac{Z_s}{Z_0} \int_\Gamma (\mathbf{n} \times \mathbf{H}) \cdot (\mathbf{n} \times f) ds \\
= (i - 1) k_0^3 \sqrt{c\mu_0 \mu_r / 2\sigma_s Z_0^2} \int_\Gamma (\mathbf{n} \times \mathbf{H}) \cdot (\mathbf{n} \times f) ds.
$$

With this, the weak form (6.7) becomes

Find $k_0 \in \mathbb{C}$ and $\mathbf{H} \in V$, $\mathbf{H} \neq \mathbf{0}$, such that for all $f \in V$ and all $q \in W$

$$
\int_\Omega \left[ \frac{1}{\varepsilon_r} \nabla \times \mathbf{H} \cdot \nabla \times f - k_0^2 \mu_r \mathbf{H} \cdot f \right] dx \\
+ (i - 1) k_0^3 \sqrt{c\mu_0 \mu_r / 2\sigma_s Z_0^2} \int_\Gamma (\mathbf{n} \times \mathbf{H}) \cdot (\mathbf{n} \times f) ds = 0, \\
\int_\Omega \mu_r \mathbf{H} \cdot \nabla q dx = 0.
$$

(6.8)

We discretize problem (6.8) with the finite element Ritz-Galerkin method [39] employing appropriate finite element subspaces of $V$ and $W$. To that end we triangulate $\Omega$ by tetrahedra. The magnetic vector functions in $V$ are then approximated by Nédélec edge elements, while the scalar functions in $W$ are approximated by Lagrange nodal finite elements [39]. This approach avoids the
6 Cavity eigenmodal analysis with surface impedance boundary conditions

generation of spurious eigensolutions, and imposing the boundary conditions is straightforward [39].

Let the vector functions $\mathbf{N}_i$, $1 \leq i \leq n$, be the Nédélec basis functions, while the scalar functions $N_\ell$, $1 \leq \ell \leq m$, denote the Lagrange basis functions. Eventually, we obtain a constrained complex nonlinear eigenvalue problem

$$
T(\lambda)\mathbf{x} = A\mathbf{x} + \lambda^{3/2} R\mathbf{x} - \lambda^2 M\mathbf{x} = 0, \quad (6.9a)
$$

$$
C^T\mathbf{x} = 0. \quad (6.9b)
$$

Here $\lambda(= k_0)$ is the eigenvalue and $\mathbf{x}$ is the eigenvector. The matrices $A$, $R$, $M$, and $C$ in (6.9) have the entries

$$
a_{ij} = \int_\Omega \varepsilon_r^{-1} (\nabla \times \mathbf{N}_i) \cdot (\nabla \times \mathbf{N}_j) \, d\mathbf{x},
$$

$$
m_{ij} = \int_\Omega \mu_r \mathbf{N}_i \cdot \mathbf{N}_j \, d\mathbf{x},
$$

$$
r_{ij} = (i - 1) \int_\Gamma \sqrt{\frac{\varepsilon_\mu_0 \mu_r}{2\sigma_s Z_0^2}} (\mathbf{n} \times \mathbf{N}_i) \cdot (\mathbf{n} \times \mathbf{N}_j) \, d\mathbf{x},
$$

$$
c_{i\ell} = \int_\Omega \mu_r \mathbf{N}_i(\mathbf{x}) \cdot \nabla N_\ell(\mathbf{x}) \, d\mathbf{x},
$$

$$
1 \leq i, j \leq n, \quad 1 \leq \ell \leq m.
$$

We solve the nonlinear eigenproblem (6.9). Then the angular frequency $\omega$ and the decay rate $\alpha$ are derived from $\lambda$. The magnetic field $\mathbf{H}$ is obtained from the calculated eigenvector, the electric field $\mathbf{E}$ by differentiation.

The system’s total stored energy $U$ in $\Omega$, and the average power loss $P_s$ in the surface conductor $\Gamma$, are computed as [63]

$$
U = \frac{\mu_0 \mu_r}{2} \int_\Omega |\mathbf{H}|^2 \, d\mathbf{x},
$$

$$
P_s = \frac{R_s}{2} \int_\Gamma |H_\ell|^2 \, d\mathbf{x}. \quad (6.10)
$$

Here, $H_\ell$ is the tangential component of the magnetic field on $\Gamma$ [63]. The systems quality factor can then be defined by

$$
Q = \frac{\omega U}{P_s}. \quad (6.11)
$$
6.2 Nonlinear eigensolver

To solve the constrained nonlinear eigenproblem

\[ T(\lambda)x = Ax + \lambda^{3/2}Rx - \lambda^2Mx = 0, \quad C^T x = 0, \quad (6.12) \]

we apply the nonlinear Jacobi–Davidson (NLJD) method, see Alg. 5. The algorithm is taken from Betcke and Voss [8, 81], with only a few modifications.

**Algorithm 5** Nonlinear Jacobi–Davidson algorithm for computing a solution of the nonlinear eigenvalue problem (6.12).

**Input:** Matrices \( A, R, M \); initial vector \( v_0 \); target \( \tau \); tolerance \( \varepsilon \).

**Output:** Eigenpair \((\lambda, u)\) that is closest to \( \tau \).

1: Construct the matrix \( T = A + \tau^{3/2}R - \tau^2M \), and its associated preconditioner \( K \approx T^{-1} \).

2: Set initial search space \( V = \{v_0\} \); \( j = 0 \).

3: Project \( v_0 \) on \( \mathcal{R}(C)^\perp \), s.t. \( C^T v_0 = 0 \).

4: while true do

5: \( j = j + 1 \).

6: Compute the eigenpair \((\lambda_j, x)\) of the projected eigenproblem \((V^*AV + \lambda^{3/2}V^*RV - \lambda^2V^*MV)x = 0\) with \( \lambda_j \) closest to \( \rho_k \).

7: Set Ritz vector \( u = Vx, u = u/\|u\| \); residual \( r = (A + \lambda_jR - \lambda_j^2M)u \).

8: if \( \|r\|_2 < \varepsilon \) then

9: accept the Ritz pair \((\lambda_j, u)\) as the eigenpair and return.

10: end if

11: Reduce the search subspace \( V \) by restart if necessary.

12: Set \( p = \frac{3}{2} \frac{1}{\lambda_j^2}Rx - 2\lambda_jMu \).

13: Find an approximate solution of the correction equation

\[ \left( I - \frac{pu^*}{u^*p} \right) T \left( I - \frac{uu^*}{u^*u} \right) t = -r, \quad t \perp u. \quad (6.13) \]

by a Krylov solver preconditioned by \( K \).

14: Project \( t \) on \( \mathcal{R}(C)^\perp \), s.t. \( C^T t = 0 \).

15: Orthogonalize \( t \) against the current search space, \( t = t - VV^*t \), and normalize \( t = t/\|t\|_2 \).

16: Expand subspace \( V = [V, t] \).

17: end while

In steps 3 and 14, we impose the divergence-free condition \( C^T x = 0 \). We construct an appropriate projector to assert that each vector in the search...
space is in the null space of $C^T$, see [3, 23]. In step 6, we use the successive linear method [80] to solve a small-sized projected eigenproblem. To (approximately) solve the correction equation (6.13), we employ the Bi-CGStab Krylov solver together with Jacobi preconditioning. For more details of the algorithm, we refer the readers to [8, 81, 80] The Alg 5 is implemented as part of the software package Femaxx-2.0 [23].

6.3 Numerical experiments

All simulations have been performed on the Cray XT6 at the Swiss National Supercomputing Centre (CSCS) [14].

6.3.1 Pillbox cavity

Before simulating the transverse deflecting cavity, we validated the correctness and reliability of our method by means of the elementary pillbox cavity. Let radius and length of the pillbox be $r = 0.05\, \text{m}$ and $h = 0.1\, \text{m}$, respectively. The mesh contains 306,337 tetrahedra. By using quadratic Nédélec elements, the finite element discretization counts 1,986,080 degrees of freedom (dof). We compute the fundamental TM010 mode with varying surface conductivity $\sigma_s$. $\sigma_s$ is large enough that the two assumptions for a good conductor are satisfied. The results we obtained are listed in Table 6.1.

Table 6.1: Numerical analysis for TM010 mode of the pillbox cavity. $f = \frac{\omega}{2\pi}$ is the resonance frequency; $\delta$ is the skin depth (6.5); $Q$ is the quality factor (6.11).

<table>
<thead>
<tr>
<th>$\sigma_s$ (S/m)</th>
<th>$f$ (GHz)</th>
<th>$\delta$ (mm)</th>
<th>$Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^2$</td>
<td>2.238100</td>
<td>1.1</td>
<td>30</td>
</tr>
<tr>
<td>$10^4$</td>
<td>2.277208</td>
<td>0.33</td>
<td>99</td>
</tr>
<tr>
<td>$10^6$</td>
<td>2.289538</td>
<td>0.11</td>
<td>316</td>
</tr>
<tr>
<td>$10^8$</td>
<td>2.293434</td>
<td>0.033</td>
<td>1001</td>
</tr>
<tr>
<td>$5.8 \cdot 10^7$</td>
<td>2.295160</td>
<td>0.0014</td>
<td>24160</td>
</tr>
<tr>
<td>$10^{10}$</td>
<td>2.295229</td>
<td>0.00011</td>
<td>317255</td>
</tr>
</tbody>
</table>

If $\sigma_s$ decreases, the field can penetrate into the surface conductor with larger skin depth $\delta$. We understand an increasing $\delta$ is equivalent to increasing the volume of the cavity $\Omega$. Therefore, the resonant wavelength increases and the
6.3 Numerical experiments

Figure 6.1: We show the electric field distribution $|\mathbf{E}|$ for TM010 mode of a pillbox cavity with surface conductivity $\sigma_s = 5.8 \cdot 10^7$ S/m, visualized in the half domain.

frequency $f$ decreases. Our numerical results clearly show this trend in $\delta$ and $f$.

We also calculated the case of perfect electric conductor (PEC) boundary conditions. Then, $\sigma_s = \infty$ and $\mathbf{n} \cdot \mathbf{E} = 0$ on $\Gamma$. The computed resonance is $f_{\text{pec}} = 2.295234$ GHz [26]. $f_{\text{pec}}$ is slightly larger than 2.295229 GHz, where $\sigma_s = 10^{10}$ S/m. We observe that if $\sigma_s$ is $5.8 \cdot 10^7$ S/m (the conductivity of copper), a further increase of $\sigma_s$ generates almost negligible shift in $f$, i.e., less than 1 MHz. Therefore, copper is a material well suited for the cavity wall. Figure 6.1 shows the electric field distribution when $\sigma_s$ is $5.8 \cdot 10^7$ S/m.

The quality factor decreases significantly as $\sigma_s$ decreases. If $\sigma_s$ reduces by a factor of 100, then $\delta$, $R_s$ and thus $P_s$ increase by a factor of about 10. Therefore, the quality factor reduces by approximately a factor 10.

6.3.2 Transverse deflecting cavity

We simulated the 5-cells transverse deflecting cavity. Each cell is a pillbox, and the middle cell is coupled to a rectangular feeding waveguide, see 1.1(a). The detailed geometrical parameters of this cavity are given in [60].

We computed the operating TM110 mode with the surface of the conductor being copper, i.e., $\sigma_s = 5.8 \cdot 10^7$ S/m. In order to maintain accuracy, we use a mesh containing 872'261 tetrahedra. By employing quadratic Nédélec elements, the finite element discretization counts 5'726'536 degrees of freedom. 512 cores are used on the Cray XT6.
6 Cavity eigenmodal analysis with surface impedance boundary conditions

We obtained the frequency \( f = 2.995488 \text{GHz} \) and the quality factor \( Q = 15504 \). They match well with the design parameters \cite{60}. The skin depth is \( \delta \approx 1.2 \mu \text{m} \), much smaller than the radius of the cells (58.28 mm). If we replace the SIBC with PEC boundary conditions, the resonance increases slightly from \( f = 2.995488 \text{GHz} \) to \( f_{\text{pec}} = 2.995930 \text{GHz} \). The electric field distribution \(|E|\) is similar to Fig. 1.1(b).
The JDQZ eigensolver and NLJD eigensolvers are implemented as part of the electromagnetic software tool Femaxx [23, 3, 29], which is continuously being developed at the Swiss Federal Institute of Technology (ETH) Zurich and the Paul Scherrer Institute (PSI). Femaxx targets solving various kinds of electromagnetic eigenmodal problems based on the finite element method in 3-dimensional space. In order to solve very large scale problems, the code has been efficiently parallelized to profit from today’s widespread, distributed memory parallel computers. MPI (message passing interface) is the communication standard between processors.

7.1 Mesh

We use the finite element method to meet the requirements of modern electromagnetic structures which generally exhibit complicated and delicate geometric features. We use the finite element mesh generator Gmsh [24] to define the computational domain Ω which is then tessellated by Gmsh into a tetrahedral mesh. Inside the computational domain Ω internal boundaries are allowed. Gmsh will ensure that the tetrahedra respect these interfaces. Therefore, Ω can contain different material regions [24]. The mesh contains information on each entity (node, edge, face or tetrahedron), including

1. unique global ID for each entity: the global finite element assembly is based on these global indices.

2. material ID for each tetrahedron: then each tetrahedron element obtains the associated material properties.

3. surface ID for each face on the boundary: then each boundary face obtains the associated boundary conditions.

A suitable mesh data distribution and management reduces communication cost and balances the computational load. In Femaxx, we use the graph par-
7 Implementation

ParMETIS \[59\] for load balancing and parallel distribution of mesh data. ParMETIS distributes the mesh in such a way that the number of mesh nodes per core is balanced \[3, 59\]; thus, approximately balances the degrees of freedom which relate to tetrahedral edges and faces. ParMETIS also minimizes the number of edge cuts, which minimizes the communication overhead by concentrating elements in diagonal blocks and reducing the number of nonzero off-diagonal blocks \[3, 59\]. Femaxx then defines a graph which contains tetrahedral connectivity information for each node, edge, and face of the finite element mesh \[3\] in order to balance the computational load in an optimum way.

7.2 Trilinos

The Trilinos Project is an ongoing effort to develop parallel numerical linear algebra algorithms and libraries within an object-oriented software framework \[35, 75\]. It solves large-scale, complex multi-physics engineering and scientific applications. It is primarily written in C++ and provides interfaces to essential BLAS and MPI libraries.

The implementation of Femaxx is based on the Trilinos (version 10). Several packages of Trilinos are used, including Epetra, ML, Amesos and AztecOO \[3\]. For a detailed overview of the Trilinos packages, we refer the reader to \[75\].

The fundamental package Epetra defines basic parallel real-valued objects such as multi-vectors and sparse matrices. It provides the basic algebraic operations, and the underlying foundation for all higher level packages. In Femaxx, the sparse matrices uses compressed row storage (CRS) format, and they are distributed row-wise. Epetra does not support complex objects. (There is a package Tpetra \[75\] which supports complex objects, but its functionality is rather limited.) Therefore, the complex data-typed JDQZ and NLJD eigensolvers employs package Epetra, by treating real and imaginary parts separately for all complex objects. This splitting strategy is advantageous if the structures of real and imaginary parts of a matrix differ. For instance, the quadratic eigenvalue problem (3.15) in chapter 3 has matrices \(A\) and \(R\). \(A\) is real matrix and \(R\) is purely-imaginary matrix. By separately storing the real and imaginary parts of \(A\) and \(R\), the memory usage has been nearly approximately halved. On the other hand, the real and imaginary parts of matrices \(M\) or \(C\) in Eq. (3.15) have the same sparsity structure, thus the splitting does not bring any benefits but no disadvantages either. We note that the implementation of the splitting strategy is much more involved. A single complex matrix-vector multiplication is decomposed into four real matrix-vector...
multiplications, e.g.,

\[ Mx = M_r x_r - M_i x_i + i(M_r x_i + M_i x_r). \]

Here, \( M_r \) and \( M_i \) are the real and imaginary parts of the matrix \( M \). \( x_r \) and \( x_i \) are the real and imaginary parts of complex-valued vector \( x \). The Eq. (3.17) shows the linearized form of the quadratic eigenvalue problem. In order to minimize the memory usage, only the individual block matrices \( A, R, M \) and \( C \) are stored, while the double sized matrices \( \mathcal{A}, \mathcal{M} \) or \( \mathcal{C} \) are not formed.

The row-wise distribution of Epetra objects is done by specifying a communicator and a map. The package Epetra provides a communicator class, which provides the basic MPI communication functions. The map relates the global and local row indices, and describes the distribution for matrices and multi-vectors among different cores. In order to perform the basic matrix-vector product \( y = Ax \), the maps of the three objects (\( y, A \) and \( x \)) must match.

### 7.3 Postprocessing

The Femaxx post-processor can evaluate the electromagnetic field at any spatial location in the computational domain. (If the sampling location is outside the computational domain, the field is evaluated as zero.) At present, there are three specific types of sample evaluation methods available

1. the field is evaluated at the center of each tetrahedron.
2. the field is evaluated on a cartesian grid.
3. the field is evaluated on a cylindrical grid.

Finally, the electromagnetic fields can be visualized in Paraview [58].

### 7.4 Object-oriented design

Femaxx is written in C++, and exploits object-oriented design pattern [45]. For instance, Femaxx contains a eigensolver abstract class, including only the virtual member functions. The JDQZ class or NLJD class inherit from this abstract class, and provide implementations for all of the abstract methods. Therefore it is easy to extend Femaxx in the future, with other types of eigensolvers.
In this final chapter, we reflect on the current state of the 3-dimensional Femaxx electromagnetic eigensolver package and we comment on possible, future directions of research. Femaxx has been developed to a level where it is capable of solving some of the most advanced and challenging problems in the analysis of plasmonic nanostructures and, generally, highly nonlinear eigenvalue problems. While we have numerically analyzed specific structures from the field of nanooptics reliably and accurately, we nevertheless comment that it is always useful to gain as much a priori information on the underlying physics of a structure. This is especially so since in any case in the eigenmodal analysis of plasmonic structures we always end up with the solution of large eigenvalue problems where the matrix elements are generally complex-valued and, thus, it is a challenge to implement and arrive at a robust, iterative Jacobi-Davidson eigensolver scheme.

In the Femaxx package, we have implemented two different schemes for plasmonic analysis, namely (a) the scheme which solves the quadratic eigenvalue problem by linearization, chapter 3, and (b) the scheme which solves the fully nonlinear eigenvalue problem, chapter 4. While, chronologically, we have implemented the linearized scheme (a) first and applied it to the analysis of optical antennas [29], we nevertheless always recommend to use the fully nonlinear scheme (b) for the analysis of plasmonic structures. There are several reasons for this recommendation: (i) scheme (a) doubles the dimension of the eigenvalue problem, and therefore also considerably increases the computational load; (ii) scheme (b) is significantly less demanding w.r.t memory and calculation times are shorter, cf. section 4.

In certain cases, in the analysis of plasmonic nanostructures, when using quadratic Nédélec elements, both schemes, (a) and (b), experience converge problems, section 3.4.3. On the other hand, we can robustly solve the eigenvalue problem using linear Nédélec elements. Then, we do not observe convergence problems and results are robust. Therefore, we suggest that one future avenue work should be along the line to improve the eigensolver’s stability and
robustness when using quadratic Nédélec elements. Experience shows that the selection of an appropriate start vector for the eigensolver is essential for convergence. For example, one could reorganize the start vector by specifically numbering the linear degrees of freedom before the quadratic ones. Then, we first solve the eigenvalue problem using only linear Nédélec elements and, in a second step, we employ the linear solution as a starting vector for the complete quadratic Nédélec elements.

Femaxx generally employs the Jacobi preconditioning scheme in the case of linear Nédélec elements; it employs the hierarchical basis preconditioner for quadratic Nédélec elements, where each block is, again, approximated by the Jacobi preconditioner. Generally these strategies are cheap. In order to further improve the eigensolver’s convergence behavior, we suggest to include several advanced preconditioning techniques, e.g. incomplete LU (ILU) decomposition, or multi-grid schemes, preferably the algebraic variant. In particular, a preconditioner library should natively operate on complex numbers and be parallelized for distributed memory computing architectures. Unfortunately, Trilinos does not presently support robust complex number packages.

Femaxx is now in the position to reliably numerically analyze fabricated plasmonic nanostructures. But, the field of practical nano-optics has much more in store; for example, there are nanostructures that support so-called Fano resonances, that arise from the interference between radiative and non-radiative modes [21]. So far we have shown first calculations with Femaxx in [28].

We also emphasize that any further development of the Femaxx package will enormously profit from the close interaction with nano-fabrication experts, since their insights have always proved fruitful for the efficient, numerical modeling.
Femaxx User Guide

This is a brief user guide for Femaxx-2.0. We show examples for analyzing several electromagnetic structures. The simulation platform is CRAY XT6 at CSCS [14]. The available eigensolvers are `femaxx_lep_cdriver`, `femaxx_qep_cdriver` and `femaxx_nlep_cdriver`, which solves the linear, quadratic and nonlinear eigenvalue problems, respectively. The postprocessing driver is named `femaxx_cpost3d`.

**A.1 femaxx_lep_cdriver**

By considering the imperfect dielectric materials in resonator cavities, we obtain a complex linear eigenvalue problem, see Section 2.4. `femaxx_lep_cdriver` solves this kind of problems, and the associated options are

- **--help**: show the manual page.
- **--mesh (= STRING)**: input mesh file in HDF5 format.
- **--material (= STRING)**: input material file. The format of the material file is discussed later. If this file is missing, `femaxx_lep_cdriver` takes the whole computational domain as vacuum, $\mu_r = \varepsilon_r = 1$.
- **--start-vector (= STRING)**: types of the start vector $v_0$:
  - ‘one’ (default): all entries of the $v_0$ are 1.
  - ‘random’: entries of the $v_0$ have random values between 0 to 1.
- **--tau-real (=DOUBLE)**: the real part of the target eigenvalue.
- **--tau-imag (=DOUBLE)**: the imaginary part of the target eigenvalue.
- **--order (=INT)**: finite element order (possible values=1 or 2, default=1).
- **--kmax (=INT)**: number of eigensolutions to be computed (default=3).
- **--tol (=DOUBLE)**: error tolerance for eigenpairs (default=$10^{-8}$).
- **--eig solver** (=STRING): default='jdqz'. Currently only the JDQZ eigen-solver is implemented.

- **--aprec** (=STRING): preconditioner type for solving correction equation. If order=1, the possible values='diag', 'lu', 'identity', default='diag'. If order=2, the possible values='diag', '2level', default='diag'.
  - 'diag': Jacobi preconditioning.
  - 'lu': preconditioner by LU decomposition.
  - 'identity': no preconditioner.
  - '2level': hierarchical basis preconditioner.

- **--a11 solver** (=STRING): solver for (1, 1) block if order=2 and '2level' preconditiner is used. Possible values='lu', 'diag', 'identity', default='diag'.

- **--a22 solver** (=STRING): solver for (2, 2) block if order=2 and '2level' preconditiner is used. Possible values='diag', 'identity', default='diag'.

- **--h solver** (=STRING): solver for $H$ matrix, default='bicgstab'. Currently only the Bi-CGStab sparse solver is implemented.

- **--htol** (=DOUBLE): the error tolerance for solving with $H$ (default=$10^{-15}$).

- **--hitmax** (=INT): maximum number of iterations for solving with $H$ (default='150').

- **--jmin** (=INT): minimum size of the search space (default='5').

- **--jmax** (=INT): maximum size of the search space (default='15').

- **--jitmax** (=INT): maximum Jacobi-Davidson outer iteration steps (default='200').

- **--jitmax1** (=INT): maximum Jacobi-Davidson outer iteration steps for computing each eigenpair (default='50').

- **--jlimitmax** (=INT): maximum inner iterations for solving the correction equation (default='50').

- **--jtol decay** (=DOUBLE): the factor by which the inner error tolerance is decreasing (default='1.2').

- **--jtol track** (=DOUBLE): the threshold for replacing the petrov pair in correction equation (default=$10^{-8}$).
• --eigendatafile (=STRING): the name of file containing computed HDF5 formatted eigendata.

Two input files are required: the mesh HDF5 file [34] and the material file. In order to apply the perfect electric conductor boundary condition, each surface element receives an boundary ID = 0 stored in the mesh file. Usually there are more than one type of materials inside the resonator structure. Let $N$ be the total number of material types, and each receives an unique material ID (between 1 to $N$) stored in the mesh file. Then the input material file records the material properties ($\mu_r$, $\varepsilon_r$ and $\sigma$) for every material type, respectively.

\[
\begin{array}{ccccccc}
N & \text{Re}(\mu_1^1) & \text{Im}(\mu_1^1) & \text{Re}(\varepsilon_1^1) & \text{Im}(\varepsilon_1^1) & \text{Re}(\sigma_1^1) & \text{Im}(\sigma_1^1) \\
\text{Re}(\mu_2^2) & \text{Im}(\mu_2^2) & \text{Re}(\varepsilon_2^2) & \text{Im}(\varepsilon_2^2) & \text{Re}(\sigma_2^2) & \text{Im}(\sigma_2^2) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\text{Re}(\mu_N^N) & \text{Im}(\mu_N^N) & \text{Re}(\varepsilon_N^N) & \text{Im}(\varepsilon_N^N) & \text{Re}(\sigma_N^N) & \text{Im}(\sigma_N^N)
\end{array}
\]

For instance, the lossy dielectric block cavity (in Section 2.4) contains two types of material: the lossless region ($\mu_r = \varepsilon_r = 1.0$, $\sigma = 0$), and the lossy dielectric block ($\mu_r = 1.0$, $\varepsilon_r = 10 - 2i$, $\sigma = 0$). The corresponding material file (named mat.dat) is

\[
\begin{aligned}
2 \\
1.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & \%\% \text{ lossless part of the cavity} \\
1.0 & 0.0 & 10.0 & -2.0 & 0.0 & 0.0 & \%\% \text{ imperfect dielectric block}
\end{aligned}
\]

We show a batch script submitted to CRAY XE6 at CSCS.

```
#!/bin/bash
#SBATCH --ntasks=32
#SBATCH --ntasks-per-node=32
#SBATCH --time=00:10:00
set -ex
cd $SCRATCH/femaxx-meshes/lossyblockcavity/
export MPICH_ABORT_ON_ERROR=1
export MPICH_PTL_UNEX_EVENTS=40960
aprun -n 32 femaxx_lep_cdriver --mesh=lossyblockcavity.h5
--tau-real=1.0 --tau-imag=0.0 --jitmax=300 --jitmax1=50 --kmax=6
--order=1 --tol=1.0e-8 --eigsolver=jdqz --aprec=diag --hprec=diag
--htol=1e-15 --hitmax=200 --jtoldecay=1.2 --jmax=25 --jmin=10
--jlinitmax=100 --material=mat.dat
--eigendatafile=lossyblockcavity_output.h5
```
exit

Here ‘aprun’ is the job launcher for Cray XE6 at CSCS [14]

A.2 femaxx_qep_cdriver

femaxx_qep_cdriver solves the complex quadratic eigenvalue problem. It is useful in several applications:

1. ohmic lossy conductors, see Section 2.5.
2. DRA, see Section 3.4.1.
3. plasmonic nanostructures, see Chapter 3.

The femaxx_qep_cdriver options are

- **--help**: show the manual page.
- **--mesh (=STRING)**: input mesh file in HDF5 format.
- **--material (=STRING)**: input material file.
- **--start-vector (=STRING)**: types of the start vector $v_0$:
  - ‘one’ (default): all entries of the $v_0$ are 1.
  - ‘random’: entries of the $v_0$ have random values between 0 to 1.
  - ‘nano’: start vector designed for application 3, see Section 6.2.
- **--tau-real (=DOUBLE)**: the real part of the target eigenvalue.
- **--tau-imag (=DOUBLE)**: the imaginary part of the target eigenvalue.
- **--order (=INT)**: finite element order (possible values=1 or 2, default=1).
- **--kmax (=INT)**: number of eigensolutions to be computed (default=3).
- **--tol (=DOUBLE)**: error tolerance for eigenpairs (default=$10^{-6}$).
- **--eigsolver (=STRING)**: default=’jdqz’. Currently only the JDQZ eigensolver is implemented.
- **--blkprec (=STRING)**: preconditioner type for solving correction equation. The possible types are
- ‘blkGS’ (default): $2 \times 2$ block Gauss-Seidel preconditioner.
- ‘diag’: Jacobi preconditioning.
- ‘identity’: no preconditioner.

- --blk11prec (=STRING): preconditioner type for solving the (1,1) block. The possible types are ‘diag’, ‘lu’, default=‘diag’.

- --hsolver (=STRING): solver for $H$ matrix, default=‘bicgstab’. Currently only the Bi-CGStab sparse solver is implemented.

- --htol (=DOUBLE): the error tolerance for solving with $H$ (default=$10^{-15}$).

- --hitmax (=INT): maximum number of iterations for solving with $H$ (default=’150’).

- --jmin (=INT): minimum size of the search space (default=’5’).

- --jmax (=INT): maximum size of the search space (default=’15’).

- --jitmax (=INT): maximum Jacobi-Davidson outer iteration steps (default=’200’).

- --jitmax1 (=INT): maximum Jacobi-Davidson outer iteration steps for computing each eigenpair (default=’50’).

- --jlininitmax (=INT): maximum inner iterations for solving the correction equation (default=’50’).

- --jtoldecay (=DOUBLE): the factor by which the inner error tolerance is decreasing (default=’1.2’).

- --jtoltrack (=DOUBLE): the threshold for replacing the petrov pair in correction equation (default=’$10^{-6}$’).

- --eigendatafile (=STRING): the name of file containing computed HDF5 formatted eigendata.

- --character-length (=STRING): Unit of length. The possible values are ‘meter’, ‘decimeter’, ‘centimeter’, ‘millimeter’, ‘micrometer’, ‘10nanometer’, ‘nanometer’, default=’meter’. The unit ‘10nanometer’ is $10^{-8}$ m, it is useful in application 3.

- --dispersive-material (=STRING): the possible dispersive material models are
A Femaxx User Guide

- ‘gold_DL’: Drude-Lorentz model for gold [78].
- ‘gold_exp’: experimental data of gold [40].
- ‘silver_exp’: experimental data of silver [40].

This option is required only for application 3.

Here most options are the same as in femaxx_lep_cdriver. However, since femaxx_qep_cdriver linearize the quadratic eigenproblem and yields an \(2 \times 2\) blocked linear eigenproblem, it is reasonable to use the blocked-type preconditioning \(blk11prec = \text{‘blkGS’}\).

First we show an example batch script for application 1, i.e., the half-filled rectangular resonator, cf. Section 2.5).

```bash
#! /bin/bash
#SBATCH --ntasks=32
#SBATCH --ntasks-per-node=32
#SBATCH --time=00:03:00
set -ex
cd $SCRATCH/femaxx-meshes/half-filled-rectangular-resonator/
export MPICH_ABORT_ON_ERROR=1
export MPICH_PTL_UNEX_EVENTS=40960
aprun -n 32 femaxx_qep_cdriver
  --mesh=half-filled-rectangular-resonator.h5
  --tau-real=100 --tau-imag=0 --jitmax1=60 --kmax=1 --order=1
  --tol=1.0e-7 --eigsolver=jdqz --blkprec=blkGS --blk11prec=diag
  --jtoldecay=1.2 --jtoltrack=1e-6 --jlinitmax=100
  --material=mat.dat --character-length=meter
  --eigendatafile=half-filled-rectangular-resonator_output.h5
exit
```

The input material file mat.dat is

```
2
1.0 0.0 1.0 0.0 0.0 0.0 %% lossless half of the resonator
1.0 0.0 2.0 0.0 1.3 0.0 %% the other half with ohmic loss
```

One half of the resonator is lossless (\(\mu_r = \varepsilon_r = 1, \sigma = 0\)), and the other half presents the ohmic conductivity loss (\(\mu_r = 1, \varepsilon_r = 2, \sigma = 1.3 \text{ S/m}\)).

Second we show a batch script for application 2, the DRA simulation.
The input material file mat.dat is

\[
\begin{array}{cccccc}
1.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\
1.0 & 0.0 & 37.84 & 0.0 & 0.0 & 0.0
\end{array}
\%
\text{ vacuum}
\]

\[
\begin{array}{cccccc}
1.0 & 0.0 & 37.84 & 0.0 & 0.0 & 0.0 \\
1.0 & 0.0 & 37.84 & 0.0 & 0.0 & 0.0
\end{array}
\%
\text{ DRA structure}
\]

The whole computational domain can be divided into two parts: the DRA structure (\(\mu_r = 1, \varepsilon_r = 37.84, \sigma = 0\)), and its surrounding vacuum (\(\mu_r = \varepsilon_r = 1, \sigma = 0\)). Here since we the use 1st order absorbing boundary condition (ABC), we let boundary ID be 4 (stored in DRA.h5) on all surface elements.

Finally, we display an example script for the analysis of the optical dipole antenna, cf. Section 3.4.4. (It is nevertheless not recommended to use femaxx_qep_cdriver for application 3, cf. discussion in Chapter 8. The users are suggested to use femaxx_nlep_cdriver, see below.)
A Femaxx User Guide

aprun -n 64 $HOME/extlib/femaxx-2.0/bin/femaxx_qep_cdriver
--mesh=opticaldipoleantenna.h5 --tau-real=0.07 --tau-imag=0.0
--jitmax1=50 --kmax=1 --order=1 --tol=8e-4 --eigsolver=jdqz
--blkprec=blkGS --blk11prec=diag --hitmax=200 --jmax=40
--material=mat.dat --dispersive-material=gold_exp
--character-length=10nanometer --start-vector=nano
--eigendatafile=opticaldipoleantenna.h5_output.h5

exit

Recall that we have to solve a sequence of quadratic eigenproblems, thus the material file mat.dat must be consequently updated according to the computed eigenvalue estimate Section 3.4.1 Eventually, the input material file mat.dat is

2
1.0 0.0 1.0 0.0 0.0 0.0 %% vacuum
1.0 0.0 -13.6482 1.0352 0.0 0.0 %% data for gold at 1.88eV

Here, the surrounding media is vacuum. By using the experimental data for gold (--dispersive-material = gold_exp), the two gold dipoles have \( \varepsilon_r = -13.6482 + 1.0352i \) at 1.88 eV (or 660 nm). The boundary ID is again 4 for all surface elements, on which the 1st order ABC is used.

In order to analyze plasmonic nanostructures with femaxx.qep.cdriver, the users must also follow two rules:

1. set --start-vector = ‘nano’.

2. write the dispersive material data in the last line in the material file. In other words, the dispersive material should receives the biggest material ID in the mesh HDF5 file.

A.3 femaxx.nlep.cdriver

femaxx.nlep.cdriver solves the complex nonlinear eigenvalue problem. It is capable for any quadratic eigenproblems that femaxx.qep.cdriver solves. Moreover, it also solves the ‘truly’ nonlinear eigenvalue problem. There are two major applications:

1. plasmonic nanostructures, see Chapter 4

2. RF cavity with finite surface conductivity, see Chapter 6.

The femaxx.nlep.cdriver options are:
A.3 femaxx_nlep_cdriver

- **--help**: show the manual page.
- **--mesh (= STRING)**: input mesh file in HDF5 format.
- **--material (= STRING)**: input material file.
- **--start-vector (= STRING)**: types of the start vector $\mathbf{v}_0$:
  - ‘one’: all entries of the $\mathbf{v}_0$ are 1.
  - ‘random’: entries of the $\mathbf{v}_0$ have random values between 0 to 1.
  - ‘nano’ (default): start vector designed for application 3, see Section 6.2.
- **--eigsolver (=STRING)**: Types of eigensolvers:
- **--tau-real (=DOUBLE)**: the real part of the target eigenvalue. This option is used if and only if **--eigsolver** = ‘nljdsibc’.
- **--tau-imag (=DOUBLE)**: the imaginary part of the target eigenvalue. This option is used if and only if **--eigsolver** = ‘nljdsibc’.
- **--sigma-surface (=DOUBLE)**: the surface conductivity $\sigma_s$ (default=5.8×10$^7$). This option is used if and only if **--eigsolver** = ‘nljdsibc’.
- **--sample-start (= DOUBLE)**: the start sampling point (Unit: eV). This option is used if and only if **--eigsolver** = ‘nljdnano’.
- **--sample-interval (= DOUBLE)**: the sampling interval (Unit: eV). This option is used if and only if **--eigsolver** = ‘nljdnano’.
- **--damping (= DOUBLE)**: the damping term (between 0 to 1). This option is used if and only if **--eigsolver** = ‘nljdnano’.
- **--dispersive-material (=STRING)**: the possible dispersive material models are
  - ‘gold_DL’: Drude-Lorentz model for gold [78].
  - ‘gold_exp’: experimental data of gold [40].
  - ‘silver_exp’: experimental data of silver [40].
This option is used if and only if **--eigsolver** = ‘nljdnano’.
• **--eigmatchtol-real** (=DOUBLE): the error tolerance for the computed and estimate eigenvalue’s real part (default=0.01). This option is used if and only if **--eigsolver** = ‘nljdnano’.

• **--eigtol-real** (=DOUBLE): the tolerance for the change in eigenvalue’s real part (default=0.01). This option is used if and only if **--eigsolver** = ‘nljdnano’.

• **--eigtol-imag** (=DOUBLE): the tolerance for the change in eigenvalue’s imaginary part (default=0.02). This option is used if and only if **--eigsolver** = ‘nljdnano’.

• **--extraction** (=STRING): Strategies for solving the projected eigenproblem and extract the Ritz pair: successive linear method (‘sl’), inverse iteration (‘iit’), linearizing the quadratic EP (‘lqep’) (default = ‘lqep’). If **--eigsolver** = ‘nljdnano’, all strategies are available; If **--eigsolver** = ‘nljdsibc’, only ‘sl’ and ‘iit’ are available.

• **--order** (=INT): finite element order (possible values=1 or 2, default=1).

• **--kmax** (=INT): number of eigensolutions to be computed (default=3).

• **--tol** (=DOUBLE): error tolerance for eigenpairs (default=10⁻⁶).

• **--tprec** (=STRING): preconditioner type for solving correction equation. If **order**=1, the possible values=‘diag’, ‘lu’, ‘identity’, default=‘diag’. If **order**=2, the possible values=‘diag’, ‘2level’, default=‘diag’.
  – ‘diag’: Jacobi preconditioning.
  – ‘lu’: preconditioner by LU decomposition.
  – ‘identity’: no preconditioner.
  – ‘2level’: hierarchical basis preconditioner.

• **--t11solver** (=STRING): solver for (1, 1) block if **order**=2 and ‘2level’ preconditiner is used. Possible values=‘lu’, ‘diag’, ‘identity’, default=‘diag’.

• **--t22solver** (=STRING): solver for (2, 2) block if **order**=2 and ‘2level’ preconditiner is used. Possible values=‘diag’, ‘identity’, default=‘diag’.

• **--hsolver** (=STRING): solver for $H$ matrix, default=‘bicgstab’. Currently only the Bi-CGStab sparse solver is implemented.

• **--htol** (=DOUBLE): the error tolerance for solving with $H$ (default=10⁻¹⁰).
• **--hitmax** (=INT): maximum number of iterations for solving with $H$ (default='100').

• **--jmin** (=INT): “the dimension of the search space $\geq jmin$” is a necessary (but not sufficient) condition for the convergence (default='1').

• **--jmax** (=INT): maximum size of the search space (default='25').

• **--jitmax** (=INT): maximum Jacobi-Davidson outer iteration steps (default='500').

• **--jitmax1** (=INT): maximum Jacobi-Davidson outer iteration steps for computing each eigenpair (default='100').

• **--jlinitmax** (=INT): maximum inner iterations for solving the correction equation (default='50').

• **--jtoldecay** (=DOUBLE): the factor by which the inner error tolerance is decreasing (default='1.2').

• **--jtoltrack** (=DOUBLE): the threshold for replacing the Ritz pair in correction equation (default=$10^{-6}$).

• **--eigendatafile** (=STRING): the name of file containing computed HDF5 formatted eigendata.


We employ **femaxx_nlep_cdriver** to analyze the optical antenna in Section 4.4. An example script file is

```bash
#!/bin/bash
#SBATCH --ntasks=64
#SBATCH --ntasks-per-node=32
#SBATCH --time=00:05:00

set -ex
cd $SCRATCH/femaxx-meshes/opticaldimerantenna/
export MPICH_ABORT_ON_ERROR=1
export MPICH_PTL_UNEX_EVENTS=40960

aprun -n 64 $HOME/extlib/femaxx-2.0/bin/femaxx_nlep_cdriver
```
The eigensolver ‘nljdnano’ run through the 4 sampling points \((-kmax = 4\)), with the first sampling point = 0.71 eV. The material file \texttt{mat.dat} is

\begin{verbatim}
2
1.0 0.0 1.0 0.0 0.0 0.0  \%\% vacuum
1.0 0.0 -132.9874 14.0271 0.0 0.0  \%\% data for gold at 0.75eV
\end{verbatim}

Here we use the gold data at 0.75 eV (i.e., \(\varepsilon_r = -132.9874 + 14.0271i\)), instead of 0.71 eV. In other words, even the users do not know the exact value of \(\varepsilon_r\) at a certain energy level (e.g., 0.71 eV), it is flexible to use the data at a nearby energy level (e.g., 0.75/eV) to approximate it. During iterations \texttt{femaxx_nlep_cdriver} is able to fix such initial mismatch automatically.

All the previous examples in this user guide contain 2 material types in the input \texttt{material} file. However, we emphasize that all \texttt{femaxx} solvers are flexible enough with almost arbitrary material arrangements. For example in Chapter 5 (Molecular nanosensor), we consider effects of the substrate and molecule layer. Then the \texttt{material} file is

\begin{verbatim}
4
1.0 0.0 1.0 0.0 0.0 0.0  \%\% vacuum
1.0 0.0 1.69 0.0 0.0 0.0  \%\% molecule layer
1.0 0.0 2.25 0.0 0.0 0.0  \%\% substrate
1.0 0.0 -32.041 1.9254 0.0 0.0  \%\% data for gold at 1.39eV
\end{verbatim}

There are 4 types of materials, and we let \(\varepsilon_r = 1.3^2 = 1.69\) for the molecule layer, and \(\varepsilon_r = 1.5^2 = 2.25\) for the substrate.

In the end, we show a script file for simulating the pillbox cavity (Section 6.3.1): The surface impedance boundary condition is used, and we let the boundary ID = 5 on all surface elements.

\begin{verbatim}
#!/bin/bash
#SBATCH --ntasks=128
#SBATCH --ntasks-per-node=32
\end{verbatim}
A.4 femaxx_cpost

All the three eigensolvers generate an output file in H5 format. femaxx_cpost is the postprocessing driver, it reads the H5 file and generates a VTK-formated file, which can be visualized in Paraview [58]. An example script is

```bash
#!/bin/bash
#SBATCH --ntasks=1
#SBATCH --time=00:06:00

set -ex
cd $SCRATCH/femaxx-meshes/opticaldimerantenna/
export MPICH_ABORT_ON_ERROR=1
export MPICH_PTL_UNEX_EVENTS=40960

aprun -n 1 $HOME/extlib/femaxx-2.0/bin/femaxx_cpost3d
--input-file=opticaldimerantenna_output.h5
--output-file=opticaldimerantenna_output

Note that femaxx_cpost is not parallelized, so only 1 core is used.
The finite element method (FEM) has enjoyed wide popularity and is a major numerical tool for simulation of electromagnetic structures, see the textbooks [39, 79, 51]. In this appendix we briefly introduce the finite element method to establish the general formulation of the electromagnetic eigenvalue problems. We use the example, ohmic conductivity loss in volume regions, in Section 2.5 to demonstrate the basic steps of FEM. The associated time-harmonic electric field curl-curl equation is

\[ \nabla \times (\mu_r^{-1} \nabla \times E(x)) + ik_0 \sigma Z_0 E(x) - k_0^2 \varepsilon_r E(x) = 0, \quad x \in \Omega, \]  

(B.1)

with the divergence-free condition

\[ \nabla \cdot (\varepsilon_r E(x)) = 0, \quad x \in \Omega. \]  

(B.2)

Here, \( \Omega \subset \mathbb{R}^3 \) is a connected bounded domain, \( k_0 \) is the complex wavenumber in free space; \( \mu_r, \varepsilon_r \) are relative magnetic permeability and relative electric permittivity, respectively; \( \sigma \) is the volume ohmic conductivity. \( Z_0(=\sqrt{\mu_0/\varepsilon_0} \approx 377\Omega) \) is the characteristic impedance of free space. On the boundary \( \Gamma = \partial \Omega \) of \( \Omega \), the perfect electric conductor (PEC) boundary condition applies,

\[ \mathbf{n} \times E(x) = 0, \quad x \in \Gamma. \]  

(B.3)

In order to apply the finite element method we need a weak form of (B.1) and (B.2). Before deriving the weak form we recall a tool from vector analysis. We first note that for sufficiently smooth vector functions \( E(x) \) and \( f(x) \) that satisfy the boundary condition (B.3) we have [39, 79]

\[ \int_{\Omega} f(x) \cdot (\nabla \times \mu_r^{-1} \nabla \times E(x)) \, dx = \]  

\[ \int_{\Omega} \mu_r^{-1} (\nabla \times f(x)) \cdot (\nabla \times E(x)) \, dx - \int_{\Gamma} \mu_r^{-1} \mathbf{n} \cdot (f(x) \times \nabla \times E(x)) \, dx. \]  

(B.4)
Here we have used the so-called *First vector Green’s theorem* [63]. \( \mathbf{n} \) denotes the unit vector normal to the boundary \( \Gamma \) pointing outward. By using the vector identity [63]

\[
\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}),
\]

and the PEC boundary condition (B.3)

\[
\mathbf{n} \times \mathbf{E}(\mathbf{x}) = 0, \quad \mathbf{n} \times \mathbf{f}(\mathbf{x}) = 0, \quad \mathbf{x} \in \Gamma,
\]

the surface integral in (B.4) vanishes:

\[
\int_{\Gamma} \mu^{-1} \mathbf{n} \cdot (\mathbf{f}(\mathbf{x}) \times \nabla \times \mathbf{E}(\mathbf{x})) \, d\mathbf{x} = \int_{\Gamma} \mu^{-1} (\nabla \times \mathbf{E}(\mathbf{x})) \cdot (\mathbf{n} \times \mathbf{f}(\mathbf{x})) \, d\mathbf{x} = 0.
\]

Therefore the Eq. (B.4) becomes

\[
\int_{\Omega} \mathbf{f}(\mathbf{x}) \cdot (\nabla \times \mu^{-1} \nabla \times \mathbf{E}(\mathbf{x})) \, d\mathbf{x} = \int_{\Omega} \mu^{-1} (\nabla \times \mathbf{f}(\mathbf{x})) \cdot (\nabla \times \mathbf{E}(\mathbf{x})) \, d\mathbf{x}. \quad (B.5)
\]

Now we turn to the weak form of (B.1). We multiply the Eq. (B.1) with a so-called test function \( \mathbf{f}(\mathbf{x}) \) from both sides and then integrate over \( \Omega \). By using Eq. (B.5) we obtain

\[
\int_{\Omega} \mu^{-1} (\nabla \times \mathbf{f}) \cdot (\nabla \times \mathbf{E}) \, d\mathbf{x} + ik_0 \int_{\Omega} \sigma Z_0 \mathbf{f} \cdot \mathbf{E} \, d\mathbf{x} + k_0^2 \int_{\Omega} \varepsilon_r \mathbf{f} \cdot \mathbf{E} \, d\mathbf{x} = 0. \quad (B.6)
\]

Similarly, we derive the weak form of divergence-free condition (B.2). For any sufficiently smooth scalar function \( q(\mathbf{x}) \) vanishing on \( \Gamma \) (i.e. \( q(\mathbf{x}) = 0, \mathbf{x} \in \Gamma \)), we have [23]

\[
\int_{\Omega} \nabla \cdot (\varepsilon_r \mathbf{E}(\mathbf{x})) \, q(\mathbf{x}) \, d\mathbf{x} = -\int_{\Omega} \varepsilon_r \mathbf{E}(\mathbf{x}) \cdot (\nabla q(\mathbf{x})) \, d\mathbf{x}. \quad (B.7)
\]

Combining Eqs. (B.6) and (B.7), a natural weak form of Eqs. (B.1) to (B.3) is:

*Find \( k_0 \in \mathbb{C} \) and \( \mathbf{E} \in V, \mathbf{E} \neq 0 \), such that for all \( \mathbf{f} \in V \) and all \( q \in W \)

\[
\int_{\Omega} \mu^{-1} \nabla \times \mathbf{f} \cdot \nabla \times \mathbf{E} \, d\mathbf{x} + ik_0 \int_{\Omega} \sigma Z_0 \mathbf{f} \cdot \mathbf{E} \, d\mathbf{x} + k_0^2 \int_{\Omega} \varepsilon_r \mathbf{f} \cdot \mathbf{E} \, d\mathbf{x} = 0,
\]

\[
\int_{\Omega} \varepsilon_r \mathbf{E} \cdot \nabla q \, d\mathbf{x} = 0. \quad (B.8)
\]
Here, $V$ denotes the functions in $H(\text{curl}; \Omega)$ that satisfy the boundary condition (B.3) and $W = H^1_0(\Omega)$. For details on these function spaces see [51]. We discretize (B.8) by employing appropriate finite element subspaces of $V$ and $W$. In Femaxx, the electric vector functions in $V$ are approximated by Nédélec (vector) finite elements [39, 79, 51]. Let the vector functions $N_i$, $1 \leq i \leq n$, be the Nédélec basis functions. The scalar functions in $W$ are approximated by Lagrange (nodal) finite elements [51]. Let the scalar functions $N_\ell$, $1 \leq \ell \leq m$, denote the Lagrange basis functions. The electric field in $\Omega$ is then approximated as

$$E = \sum_{i=1}^{n} N_i x_i, \quad \text{(B.9)}$$

where the $x_i$ denote the expansion coefficients of the basis. On substituting Eq. (B.9) into (B.8), the Eq. (B.8) becomes

$$\sum_{i=1}^{n} \left[ \int_{\Omega} \mu_r^{-1} (\nabla \times f) \cdot (\nabla \times N_i) \right] x_i + i k_0 \sum_{i=1}^{n} \left[ \int_{\Omega} \sigma Z_0 f \cdot N_i \right] x_i$$

$$+ k_0^2 \sum_{i=1}^{n} \left[ \int_{\Omega} \varepsilon_r N_j \cdot N_i \right] x_i = 0,$$

$$\sum_{i=1}^{n} \left[ \int_{\Omega} \varepsilon_r N_i \cdot \nabla q \right] x_i = 0. \quad \text{(B.10)}$$

In the Ritz-Galerkin approach, the search space is equal to the test space. Since the Eq. (B.10) holds for any vector function $f \in V$ and any scalar function $q \in W$, we let $f$ be $N_j$, $1 \leq j \leq n$ and $q$ be $N_\ell$, $1 \leq \ell \leq m$, respectively, then we have

$$\sum_{i=1}^{n} \left[ \int_{\Omega} \mu_r^{-1} (\nabla \times N_j) \cdot (\nabla \times N_i) \right] x_i + i k_0 \sum_{i=1}^{n} \left[ \int_{\Omega} \sigma Z_0 N_j \cdot N_i \right] x_i$$

$$+ k_0^2 \sum_{i=1}^{n} \left[ \int_{\Omega} \varepsilon_r N_j \cdot N_i \right] x_i = 0, \quad 1 \leq j \leq n,$$

$$\sum_{i=1}^{n} \left[ \int_{\Omega} \varepsilon_r N_i \cdot \nabla N_\ell \right] x_i = 0, \quad 1 \leq \ell \leq m. \quad \text{(B.11)}$$
B Finite Element Method

The Eq. (B.11) is in fact a constrained complex-valued quadratic eigenvalue problem. Its matrix form is

\[
T(\lambda)x = Ax + \lambda Rx - \lambda^2 Mx = 0, \\
C^T x = 0.
\]  

(B.12)

Here \( \lambda = k_0 \) is the eigenvalue and the coefficients vector \( x = [x_1, ..., x_n]^T \) is the corresponding eigenvector. According to (B.11), the matrices \( A, R, M \) and \( C \) in Eqs. (B.12) have entries

\[
a_{ij} = \int_\Omega \mu_r^{-1}(\nabla \times N_i) \cdot (\nabla \times N_j) \, dx, \quad 1 \leq i, j \leq n, \\
m_{ij} = \int_\Omega \varepsilon_r N_i \cdot N_j \, dx, \quad 1 \leq i, j \leq n, \\
r_{ij} = i \int_\Omega \sigma Z_0 N_i \cdot N_j \, dx, \quad 1 \leq i, j \leq n, \\
c_{i\ell} = \int_\Omega \varepsilon_r N_i \cdot \nabla N_\ell \, dx, \quad 1 \leq i \leq n, \; 1 \leq \ell \leq m.
\]  

(B.13)

Note that \( A, M \) and \( R \) are symmetric square matrices and \( C \) is a rectangular matrix.

In the examples in Sections 1.2.1 and 2.4, the second term (related to \( \sigma \)) of the time-harmonic electric field curl-curl equation (B.1) vanishes. Therefore, the \( R \) matrix is not generated and the quadratic eigenvalue problem (B.12) reduces to a linear eigenvalue problem, c.f. Eq (1.3).


Bibliography


Bibliography


[28] H. Guo, B. Oswald, and P. Arbenz. 3-dimensional eigenmodal analysis of plasmonic nano-structures: with a focus on molecular nanosensors, July 2012. Presentation at the 8th Workshop on Numerical Methods for Optical Nano Structures.


Bibliography


Bibliography


Bibliography


Curriculum Vitae

Hua Guo

Personal information

Date of birth       May 1st, 1984  
Place of birth     Hefei, China  
Married to         Yuanrui Zhang  
Citizenship        Chinese

Education

2008-2012           Ph.D. student and teaching assistant at the Computer Science Department at ETH Zurich  
2006-2008           MSc study (electrical engineering) at Huazhong University of Science and Technology, China  
2002-2006           BSc study (applied mathematics, physics) at Wuhan University